Methylation & genomic imprinting in the bumblebee, *Bombus terrestris*

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester

by

Crisenthiya Indunil Clayton, MSc (Liverpool)

Department of Biology

University of Leicester

December 2012

Methylation & genomic imprinting in the bumblebee,
Bombus terrestris

Crisenthiya Indunil Clayton

Abstract

Genomic imprinting, the parent-of-origin specific silencing of alleles, plays an important role in phenotypic plasticity and consequently evolution. The leading explanation for genomic imprinting is Haig's conflict theory, which suggests that alleles from each parent have evolved under different selectional pressures, resulting in the differential expression of patrigenes and matrigenes. Previous studies have mainly used mammals and flowering plants to test Haig's theory. However, there is a lack of independent evidence to support the theory.

My PhD thesis attempts to conduct an independent test of Haig's conflict theory using buff tailed bumblebee *Bombus terrestris*. A methylation system to facilitate genomic imprinting has not been found in this species. Therefore the first aim of the study was to establish the presence of a functional methylation system in *B. terrestris* before testing Haig's conflict theory using worker reproduction in queen-less colonies.

The initial finding is that a methylation system exists in *B. terrestris*. The next study, investigating the presence of methylated genes, revealed differential methylation patterns in caste and life stages. Finally, genes involved with worker reproduction in a range of social insects were identified, but distinguishing the matrigene and the patrigene for each gene was unsuccessful. Therefore the final study investigating the presence of imprinted genes in *B. terrestris* and whether they conform to the expression patterns hypothesised by Haig's conflict theory could not be analysed.

Although this study did not provide conclusive evidence to support Haig's conflict theory, the presence of methylation in genes involved with worker reproduction in reproducing and non-reproducing *B. terrestris* workers suggests that further analysis is needed. With adequate evidence, proving Haig's conflict theory will not only expand our knowledge of invertebrate methylation, but also our understanding of conflict within social insect societies and our knowledge of how genomic imprinting affects phenotypic plasticity.

Acknowledgements

I would like to thank my supervisor Dr. Eamonn Mallon for all the inspiration, guidance and advice over the years. And all the patience!

Secondly, a big thank you to past and present members of Lab 219- Carolyn, Akram, Gillian and Harindra; and members of the Leicester Biology Department, in particular Donna and Ceinwen.

My parents and two brothers, none of this would have been possible without your encouragement, confidence and support. Especially Jerome, thanks for running the ship during the last month. I would also like to thank the extended family of Claytons, Dharmawardenas, Dissanayakes, Fernandos and family friends back in Sri Lanka. Plus Rev. Fr. Chandana and the Carmelite Sisters at Katugasthota Convent for their advice and including me in their prayers.

Friends! Sanjee, thanks for believing this was possible and *very gently* nudging me along the right path from my first day as an undergraduate. Sava, it's been 16 years of madness in 2 different continents! Do you think we'll ever mean it when we say 'This is our last all-nighter ever!'? Ian, thanks for the music, trains, adventures & so much more! Gill the lab wench & most fabulous housemate ever; I swear by my pretty floral bonnet you get a double mention for making it so. Kate thanks for always being awesome! Jon, thanks for checking upon me. Ali for being a constant source of cheerfulness, the darts games & introducing me to Star Wars during thesis intervals. Rob I'm sorry I couldn't work 'Captain's Log' into a Methods section- I tried, I really did. More importantly, I'm forever grateful to all of you for not sticking me in a straitjacket when the Trekkiness reached critical levels. There's plenty more people I would like to mention, but I need to stop before I run out of space for the sciency stuff. So here's a big 'thank you' to everyone else!

Thank you God for the good days, the bad days and helping me persevere. *"I will say of the Lord. He is my refuge and my fortress: In Him will I trust."* (Psalm 91:2). Finally, thank you Mary, my spiritual Mother, for listening.

For my parents;

Thank you for the unconditional and limitless love

Ammi & Thaththi.

Table of Contents

List of Figures

List of Tables

List of abbreviations

°C……degrees centigrade
-F……forward primer
-R……..reverse primer
ml……microlitre
AIMS……..amplification of intermethylated sites
BAC……...bacterial artificial chromosome
BAH domain……bromo-adjacent homology
bp……..base pair
BSA……bovine serum albumin
cDNA……complementary DNA
CpA……cytosine-phosphate-adenine
CpG……cytosine-phosphate-guanine
dCTP……cytosine triphosphate
DMSO……dimethylsulphoxide
DNA……deoxyribonucleic acid
Dnmt1……...DNA (5'-cytosine) methyltransferase 1
Dnmt2……..DNA (5'-cytosine) methyltransferase
Dnmt3……..DNA (5'-cytosine) methyltransferase 3 subfamily
dNTPs……2‟-deoxyribose 5‟-triphosphates
EDTA……ethylenediaminetetraacetic acid
ESS……evolutionarily stable strategy
EtBr……ethidium bromide
g……grams
GO……..Gene Ontology
HL……Hyperladder
IPTG……..isopropyl-β-D-thio-galactoside
JH……juvenile hormone
kb……..kilobase
L……larvae
L1……..larval stage 1
L2……..larval stage 2
L3……..larval stage 3
L4……..larval stage 4
LB……lysogeny broth/Luria broth/Luria-Bertani medium
LDLR……...low density lipoprotein receptor
M……molar
mb……...mega base
MBD……..methyl binding domain
MBeq……..mega Becquerel's
mg……milligram
$MgCl_2$……..magnesium chloride
mRNA……...messenger RNA
mya……..million years ago
NaCl……..sodium chloride
NEB……New England Biolabs
ng……nanogram
P……pupae
P/C……phenol/chloroform

PBS……phosphate-buffered saline
PCR……...polymerase chain reaction
pMol……..pico mol
Q……..queen bee
QS1……Q-step 1 ladder
RNA……..ribose nucleic acid
Rpm…….revolutions per minute
s……...seconds
SCC...…….saline-sodium citrate
SDS…….sodium dodecyl sulphate
SNP……..……single nucleotide polymorphism
SOC…….super-optimal broth (catabolite repression)
SSCP…….single strand confirmation polymorphism
ssDNA……...single stranded DNA
T^A…….annealing temperature
TAE……..Tris-acetate-EDTA buffer
TE……..Tris-EDTA
T^E…….extension temperature
T^M……..melting temperature
V……volt
W……...worker bee
w/v ……weight per volume
µM……micromolar
X-gal……...5-bromo-4-chloro-3-indolyl-beta-D-galacto-pyranoside
YB……Yorkshire Bioscience

Chapter 1 : Introduction

1.1 Introduction

A characteristic of most organisms is the diploid nature of their genome, in that they have two alleles for each gene. However, in mammals and angiosperms (flowering plants), only one allele of the gene is expressed. In some cases the silenced allele is from the mother and in others the silenced allele is inherited from the father, leading to a parent-of-origin based silencing process called genomic imprinting. The study of imprinting is important because changes in imprinting can lead to cancer and disorders such as Prader-Willi and Angelman syndromes (Jaenisch, 1997). Genomic imprinting arose 150 million years ago (Murphy & Jirtle, 2003) and is considered an important component of evolution. However, diploidy protects against harmful recessive mutations, while genomic imprinting causes haploidy. Silencing one allele while having the option of expressing both alleles has led to suggestions that genomic imprint is an evolutionary paradox.

The leading explanation for genomic imprinting is Haig's conflict theory, which is based on the concept that alleles from each parent have evolved under different selectional pressures. In polyandrous species (where a female mates with multiple males) the patrigenes will be different in each sibling and under selectional pressure to prevent sharing resources with the organism's siblings. However, matrigenes will be selected to share resources as it is likely that a copy of a matrigene, unlike the patrigene, would be found in other siblings. Conflict has been predicted and confirmed to occur in mammals and angiosperms, as member species are directly involved with the transfer of resources from parent to offspring, via the placenta and endosperm respectively. However, an independent test of Haig's conflict theory is necessary, using a non-

mammalian or non-angiosperm test species. Social insects colonies exhibit a reproductive division of labour, where the queens reproduce and workers forgo reproduction in order to take care of the queen's young. Thus, there is a transfer of resources analogous to that of parent to offspring in mammals and plants. Social insects are also haplodiploid, with males developing from unfertilized eggs and females from fertilised eggs. Thus, different selectional pressures are present and conflicts can arise between matrigenes and patrigenes over reproduction in insect colonies. This thesis attempts to independently validate Haig's conflict theory using an example of conflict found in social insect systems, worker reproduction.

1.2 Role of phenotypic plasticity in evolutionary biology

Reproductive isolation has traditionally played an important part in speciation as it increases species diversity by producing genetically independent pedigrees. However, reproductive isolation is not solely responsible for the emergence of a new species as it cannot create phenotypic divergence. West-Eberhard (2005) argues that since selection has an effect on phenotypes as opposed to genotypes or the genes themselves, new traits can be a result of the environment. A new environment is described as a change in an organism's current environment or a result of the organism invading a new habitat (Ghalambor *et al*., 2007). A well-rounded genotype would result in a phenotype that is effective in any environment. However, this does not always occur. When an individual, population or species is subjected to a new environment, natural selection can result in the extinction of the individual, population or species (Pigliucci *et al*., 2006). Phenotypic plasticity prevents extinction by enabling multiple phenotypes to be produced by one genotype as a response to a new or changing environment (Kelly *et*

al., 2012). Consequently, plasticity allows an organism to be able to tolerate a more diverse range of environmental conditions and therefore also increases the organism's fitness over a range of environments (Pigliucci, 2001). An example of this is the production of armoured carapaces in the presence of predators, which protects a number of fish species from predation. A lack of predators results in the armour not developing (Spitze, 1992). Another example of phenotypic plasticity is how different incubation temperatures of reptilian eggs can determine the gender of the offspring (Shine, 1999). Consequently, phenotypic plasticity is a key aspect of evolutionary change.

Evolutionary ecologists classically studied phenotypic plasticity by exposing genetically related individuals such as clones, full and half siblings to different environmental conditions. The results were statistically analysed to explore genetic correlations and give predictions on the rate and direction of evolutionary change in the population (Pigliucci, 1996). However, using a solely statistical approach does not delve deeply into the genetic basis of phenotypic plasticity for any given genotype. In addition, it employs a simplistic view of numerous genes having a cumulative effect. An alternative approach is to screen for differences in DNA or protein expression in individuals subjected to different environmental conditions. For example, Santoni *et al.* (1994) compared the protein profiles for one genotype in *Arabidopsis thaliana* when the plants were subjected to normal light conditions and darkness. The results showed differential expression of proteins in an environmental or genotypic specific pattern. Using a combination of molecular and organismal approaches in multiple environment studies thus contributes to the understanding of how phenotypes evolve.

A reaction norm is the environment-dependent variation of phenotypic expression in a genotype (Storz *et al*., 2010). A genotypic variation in a reaction norm is the product of the genotype and environmental interaction (genotype x environmental interaction) (Fry, 1995). A particular reaction norm is considered adaptive if the phenotype expressed in each environmental state is parallel to the environmental-specific optimum preferred by selection (Ghalambor *et al*., 2007). When colonising a new environment, an adaptive reaction norm would shift the mean phenotype towards the new trait optimum. However, a maladaptive reaction norm would move the phenotypic mean away from the new optimum. This results in the selection of genetically based trait variation to neutralise changes in the environment (Ghalambor *et al*., 2007). If the trait optimum is stable in a newly colonised environment, a maladaptive reaction norm would need additional genetic alterations to restore the ancestral phenotype (Grether, 2005). Maladaptive plasticity can therefore form a type of 'cryptic adaptive evolution' where the phenotypic similarity between ancestral and resultant populations is due to the effect of divergent selection which counterbalanced the trait differences caused by the changing environment (Thibert-Plante & Hendry, 2010).

As stated above, phenotypic plasticity is important because of the effect plasticity can have on genetic diversity, depending on whether it is adaptive or non-adaptive plasticity. Phenotypic plasticity is considered adaptive if the plasticity increases the individual's fitness when compared to a non-plastic individual (Hughes *et al*., 2002) and therefore evolves via natural selection. Models for the evolution of adaptive phenotypic plasticity predict that phenotypic plasticity will be successful if the population is subjected to a shifting environment which produces reliable cues and if the different phenotypes produced in each environment are selected favourably. Finally,

adaptive phenotypic plasticity will be selected for if none of the resulting phenotypes demonstrates a superior fitness in all of the different environments (Moran, 1992). That is, each phenotype has to be successful in a different environment, not all environments. Nevertheless, these predictions depend on the fitness cost necessary to establish and retain the expression of phenotypic plasticity (Ernande and Dieckman, 2004) as well as the amount of gene flow between the populations found in the different environments (Sultan and Spencer, 2007).

Phenotypic plasticity is not always adaptive and can be non-adaptive. This type of phenotypic plasticity can be due to non-adaptive environmentally induced variation. Alternatively, it can be a result of biophysical or biochemical variation, such as low temperature slowing development rate in ecotherms (Newman 1992). While fitness is affected by the slow development, it is not an adaptation to the cold. Non-adaptive phenotypic plasticity has a minimal, if any effect on fitness (Hughes *et al*., 2002).

Novel environmental conditions, such as invading a new ecological niche, changes in climate or loss of habitat can cause some phenotypes to evolve differently than in the original environment (Grether, 2005). This alternative development facilitates adaptive evolution and allows the population to reach a new adaptive peak (West-Eberhard, 2003). A means by which phenotypic plasticity can occur is through heritable changes in a phenotype that arise without an alteration to the sequence of the gene responsible for the phenotype. The study of such changes is known as epigenetics. DNA is acknowledged as the building block of life and through it all components of the organism act in unison to ensure survival, with a final goal of reproduction. However,

within-locus conflicts can sometimes arise. An example of reproduction-based within-locus conflict is genomic imprinting. Genomic imprinting is the differential expression of a gene according to which parent it was inherited from. An imprint on a patrigene (a gene that is paternally derived) or a matrigene (a gene that is maternally derived) can lead to one copy of the gene being completely silenced and cause parent-specific gene expression resulting in the gene to be functionally haploid. An imprint is not necessarily inherited by the following generation (Kronauer, 2008) and is therefore a non-Mendelian epigenetic phenomenon. The next section will discuss epigenetics and genomic imprinting in detail.

1.3 Epigenetics and genomic imprinting

Epigenetics has been hailed as a way of complementing our understanding between the inert genome and the multifaceted phenotypes that are caused by multiple genes and context-based interactions (Lyko & Maleszeka, 2011). Genomic imprinting is considered an epigenetic process because the imprint is a modification to the structure of DNA instead of the DNA sequence. Imprinting is a shifting process, in that imprints must have the ability to be erased and re-established in the germline cells of each generation. The erasing and re-establishment is necessary as the imprints depend of the sex of the individual. A paternal imprint is produced if the imprint is re-established in the developing sperm, while a maternal imprint is produced if the imprint is re-established in developing oocytes (Reik, 2001).

Genomic imprinting in plant and mammalian development has been extensively documented. For example, for the *H19* gene that is conserved in mice and humans only the maternally derived allele is expressed; the paternal copy is imprinted (Ma *et al.*, 2010). Plant endosperms only express the maternally derived copy of the *FIS1/MEA* gene with the paternally derived copy being imprinted (Spielman *et al.*, 2001). Seed lethality occurs when the maternal copy is mutated. The study of epigenetic misregulations is an important field as mutation, or a loss of imprinting, leads to many human diseases. For example, the maternally derived copy of the gene for Prader-Willi syndrome on chromosome 15 is normally silent, while the paternally derived copy of the gene for Angelman syndrome on the same chromosome is normally silent. However, when the copy which should be expressed is deleted, it results in one of the two syndromes (Cassidy & Schwartz, 1998).

The existence of functionally haploid genes has been argued to provide a means of increasing the rate of evolution of genes without putting the entire genome at risk (McGowan & Martin, 1997) despite diploidy preventing harmful deleterious mutations. In addition, one of the most common mechanisms for establishing imprints, discussed later in this thesis, can cause genomic instability. There are numerous theories which attempt to explain the evolution of genomic imprinting. The ovarian time-bomb hypothesis states that genomic imprinting evolved as a method of preventing the growth of invasive trophoblast in ovary germ cell tumours in extinct placental mammals (Varmuza & Mann, 1994). Alternative theories explain genomic imprinting as a means of parthenogenesis prevention, minimising variations in expression rates and regulating chromosome additions or deletions (Hurst, 1997). These theories, however, do not comprehensively address the full extent of genomic imprinting. For example, the

ovarian time-bomb theory does not explain how seed development, where germ line tumours do not occur, can be affected by genomic imprinting. The final theory which explains the evolution of genomic imprinting is Haig's conflict theory.

1.4 Kinship theory & Haig's conflict theory

Kin selection refers to the evolutionary strategies which promote the reproductive success of an individual's relatives, even though this has a cost to the individual's own survival and reproduction (Hamilton, 1964). Therefore, kin selection occurs when a gene's performance affects the transmission of the gene to future generations, not through offspring, but via its effects on other relations who also possess the gene. The effect of kin selection depends on the degree of relatedness between the two individuals, with a higher relatedness increasing the likelihood of the allele being shared (Lizé *et al.*, 2007). The Kinship theory assumes that the genes inherited from the mother and those inherited from the father are expressed equally in the offspring. An extension of kinship theory is Haig's conflict theory of imprinting. Haig's conflict theory explains an unequal expression of alleles from each parent as being due to the genes from the mother and father being under different selectional pressures. The predictions in Haig's theory are mostly based on interactions between a mother and her offspring because this is where parent-specific selection would be strongest. Interactions between mammalian mothers and offspring are vital for development and can influence each other's fitness (Haig, 2004), as mammalian mothers provides the foetus and offspring with nutrients and protection.

Conflicts arise when matrigenes and patrigenes are not equally related to other individuals (Kronaur, 2008). In a polyandrous species the patrigenes would favour the maternal investment of the embryo. This conflict arises due to patrigenes only being present in other full siblings, but not in the mother or half-siblings. However, this additional investment leads to a cost to both the mother and maternally related half-siblings (Krauner, 2008), as it results in the mother having less resources to invest on her other offspring, as well as herself and any future offspring. Thus when analysing the role of genomic imprinting in genes concerned with growth, it has been predicted that paternally imprinted genes lead to increases in offspring growth, while maternally imprinted genes lead to decreases in growth (Haig, 1997). In the present study, a monoandrous species (where the female mates with only one male) will be used as an independent test of Haig's conflict theory and will be discussed in more detail in later chapters.

Relationships between kin can be categorized as being symmetrical or asymmetrical based on the degree of matrilineal and patrilineal kinship (Haig, 2011). Symmetric kin include the individual, direct descendants and full siblings as there is a direct degree of kinship. Parents, grandparents, aunts, uncles and cousins are considered to be asymmetric kin as they are subject to different degrees of kinship. Kinship theory advocates that genomic imprinting evolved as a means of transcriptional control at loci whose expression caused fitness consequences for asymmetric kin (Haig, 2000). Relatedness asymmetries in diploid species can be a result of dispersal of individuals due to sex bias and the mating system (Haig, 2000b). Patrilineal and matrilineal relatedness depends on the polyandry and polygyny rates, how females distribute eggs, the dispersal rate of males and females from their nests, mating sites and egg clutches

(Lizé *et al.*, 2007). Therefore, it would be expected that in polygynous species, where each male mates with more than one female, paternal genes will be under a greater selective pressure compared to maternal genes, for reducing competition amongst offspring. In polyandrous species, where each female mates with more than one male, the maternal genes would be under a greater selectional pressure compared to paternal genes. These different selectional pressures could be due to differences in paternal and maternal investment in offspring.

The conflict theory goes on to predict that evolutionary stable strategies (ESS) at a specific locus can be symmetric or asymmetric (Haig, 1997). A symmetric ESS is when an identical level of gene expression by both maternally and paternally derived alleles is expected. In this instance, matrilineal and patrilineal inclusive fitness could be decreased if expression variations are caused by mutant alleles. An asymmetric ESS occurs when different levels of expression are exhibited by either the maternally or paternally derived alleles. If the locus is not imprinted, a concession is reached- where the ESS level of expression is in-between the needs of the parental targets. However, the case changes if imprinting is present. In that case the allele which favours the higher amount is expressed and the other allele becomes silent (Haig & Westoby, 1989). This is also known as the 'loudest voice prevails principle' (Haig, 2000). Therefore if the maternally derived allele favours a higher amount of expression, the paternal allele becomes silent. When such an ESS occurs, a small rise in gene expression can cause matrilineal and patrilineal inclusive fitness to decline, while a small decrease in expression can lead to an improved patrilineal inclusive fitness and have the opposite effect on matrilineal inclusive fitness. Similarly, if the patrilineal

derived alleles favour increased levels of gene expression, the maternally derived allele is silent and the conditions above are inverted (Haig, 2000).

A number of mechanisms which lead to the epigenetic silencing of alleles have been documented. DNA is enveloped around nucleosomes composed of histone octomers, which can be affected by histone acylation and deacylation, resulting in epigenetic modifications (MacDonald, 2011). An alternative epigenetic mechanism which acts in conjunction with histone modifications is DNA methylation. Imprinting in mammals is initiated at gametogenesis and the methylation markers are maintained during embryogenesis. During this time, imprinted loci in the remainder of the genome are demethylated. Imprints in the germ line are erased and re-established according to the sex of the embryo (Riek & Walter, 2001). These processes are governed by components of the methylation system.

1.5 Methylation as an epigenetic label

Epigenetic machinery offers a supplementary level of genome regulation which results in an organism becoming flexible as a response to the environment (Suzuki & Bird, 2008). The expression of a particular matrigene or patrigene depends on the presence or absence of epigenetic labels in the gene or the gene's promoter.

Methylation is a process where a methyl group is added to the fifth carbon in the pyrimidine ring of cytosine or the sixth nitrogen of the purine ring in adenine. It has been suggested that 5 methylcytosine is the fifth nucleotide in DNA (Lister & Ecker,

2009). Independent of its acceptance as a nucleotide, the presence of 5-methylcytosine in prokaryotic, eukaryotic and mammalian DNA initiated a quest for its function. At present, methylation has been suggested to be a conserved process with diverse functions. This is best illustrated in the diversity of roles methylation has been associated with in different species of insects, such as gene regulation in the aphid *Acyrthosiphon pisum,* imprinting in the mealy bug *Planococcus citri* (Khosla *et al*., 2006) transposon regulation in the fruit fly *Drosophila melanogaster* and genome stability in the cabbage moth *Mamestra brassicae* (Field *et al.*, 2004). All of these processes are established, retained and lost via the action of methyltransferases and DNA methyl-binding proteins.

Methylation in palindromic sections of DNA has been observed in a wide variety of organisms, from the parasitic protozoan *Entamoeba hystolytica* (Fisher *et al.*, 2004), honey bee *Apis mellifera* (Wang *et al.*, 2006), malaria vector *Anopheles gambiae* (Marhold *et al.*, 2004) and the house mouse *Mus musculus* (Jackson-Grusby *et al*., 2001) to plants such as the rockcress *Arabidopsis thaliana* (Finnegan *et al.*, 1996). While both CpG (methylation tag is situated on cytosine) and CpA (methylation tag is situated on adenine) methylation can occur, higher eukaryotes have been observed to exclusively possess methylated cytosines, where the methyl group is situated at position five of the pyramidine ring (Pennings *et al.*, 2005). The distribution of these methylated cytosines vary and range from 0-3% in insects, 6% in mammals and birds to 10% in fish and amphibians and 30% in plants (Field *et al.*, 2006). The process of DNA methylation is governed by methyltransferases, a group of enzymes that act on S-adenosyl-L-methonione (AdoMet) to modify DNA at specific sites (Jeltsch, 2002) and methyl binding proteins (MBP).

The precise effect of methylation is a much-debated topic (Field *et al.*, 2004). If a complete understanding of the role of methylation is to be achieved, the effects of methylation in a single context using multiple species must be explored. However, current literature has focused on the different functions of DNA methylation according to model organism and the experimental context it is examined under. Superficial similarities in developmental defects have been suggested to portray conserved functions of DNA methylation (Field *et al.*, 2004). However, the molecular consequences of methylated genomic DNA are quite diverse. For example, loss of DNA methylation has been documented to effect apoptosis in mice (Jackson-Grusby *et al.*, 2001) and the African clawed frog *Xenopus laevis* (Stancheva *et al.*, 2001), X chromosome inactivation and chromosome stability in *M. musculus* (Panning & Jeanisch, 1996; Gaudet *et al.*, 2003) and chromosome organisation in *A. thaliana* (Soppe *et al.*, 2002). Hence, the molecular consequences would vary on the nature of epigenetic signals. In addition, it has been suggested that DNA methylation has a role in regulating global gene activity (Field *et al.*, 2004). Another possible explanation for the presence of methylation in higher organisms revolves around the evolution of vertebrates from invertebrates, which was accompanied by an increase in gene number (Field *et al.*, 2004). It has been suggested that this increase in gene number was achievable due to DNA methylation inhibiting transcriptional background noise (Bird, 1995). Supporting this theory is evidence that the change from fractional methylation (where large regions of the genome are methylated while other regions are unmethylated) to global methylation of genomes (where almost every CpG site is methylated) occurred near the supposed emergence of vertebrates (Tweedie *et al.*, 1997).

The relationship between the number of genes in a genome and DNA methylation patterns do not necessarily confirm causation. Another hypothesis for the principal role of methylation in vertebrates is that it inhibits parasitic sequence elements (Walsh & Bestor, 1999). Accordingly, the control of gene expression has been suggested to be a secondary function, although doubts have been cast regarding the existence of an association between methylation and gene control in vertebrates (Walsh & Bestor, 1999). It has also been suggested that methylation is a means of "memorising" gene activity profiles by stabilising genes silenced via alternative mechanisms (Bird, 2002).

Regardless of how methylation evolved, it has been demonstrated that methylation patterns which are specific to certain genes are inversely correlated with the gene's activity and that activation of a gene occurs if it is artificially demethylated, while *in vitro* methylation results in a repression in the activity of the gene (Razin & Cedar, 1991). Both of these observations have led to the proposal that CpG methylation is involved in gene silencing (Razin, 1998). The mechanism by which methylation could interfere with transcription has been suggested to be associated with the prevention of binding basal transcriptional machinery and other transcription factors to promoters (Kass *et al.,* 1997). Conversely, transcriptional inhibitors have been suggested to recognise methylated CpGs and turn off transcription, either by itself or together with other elements in chromatin (Kass *et al.,* 1997). According to Razin (1998), CpG methylation causes histone deacetylation, chromatin modifications and gene silencing via the formation of a transcription repressor complex, as well as the two histone deacetylases, HDAC1 and HDAC2, around mSin3A, which is another repressor. The

repression complex is found on the methyl-binding protein MeCP2, which, in addition to its methyl-binding domain (MBD), contains a transcriptional repressor domain (TRD) (Razin, 1998). This TRD shares an overlap with an area of the genome that is affected by mSin3A (Razin, 1998).

DNA methylation in prokaryotes has been suggested to be involved in the differentiation of self and non-self DNA (Jeltsch, 2002). The distinction between self and non-self cells by the methyltransferases in restriction/modification systems (RM systems) in bacteria is an important method by which bacteriophage infection is deterred. In addition, DNA methylation has also been associated with DNA replication and cell cycle control and in managing past-replicative mismatch pairs in prokaryotes (Jeltsch, 2002). Cytosine methylation has been illustrated to be concerned with not only genomic imprinting, but also with stabilising genomes and in the permanent silencing of transposons and endogenous retroviruses (Bird, 2002). Evidence illustrates that methylated cytosines prevent the expression of genes coded in transposons, prevent DNA rearrangements brought about by transposons and silence transcription between transposons and host genes (Field *et al.*, 2004).

The importance of methylation during embryonic development is exhibited in the analogous observations made in *X. laevis* embryos and in plants, where loss of methylation caused developmental defects (Stancheva & Meehan, 2000; Finnegan *et al.*, 1996). Generally, gene silencing has been associated with the presence of methylated DNA in and the area surrounding promoter regions. If this methylation is lost, transcription of the relevant gene occurs (Field *et al.*, 2004). It has been observed

that methylation patterns in vertebrates, where 5-methylcytosine occurs in CpG doublets, appear *de novo* simultaneously with development (Field *et al.*, 2004). These patterns are either upheld or discontinued by the action of methyltransferases and demethyltransferases (Wolffe *et al.*, 1999), in addition to proteins that bind to methylated DNA, histone modification units and chromatin remodelling machinery.

1.6 Methylation systems

Dnmt1, Dnmt2 and Dnmt3 are amongst the most commonly found methyltransferase proteins in eukaryotes. Dnmt1 proteins consist of a C-terminal catalytic domain that shows high levels of conservation with bacterial methyltransferases and a large N terminal, regulatory domain that initiates protein interaction (Schaefer & Lyko, 2007). As such, it has been suggested that this characteristic of Dnmt1 enables it to interact with the components of DNA replication. Additionally, since this group of enzymes act upon hemimethylated DNA (Yoeder *et al.*, 1997), it has been proposed that Dnmt1 is responsible for the maintenance of methylation from parental to daughter strand (Schaefer & Lyko, 2007). It has been suggested that hemimethylated DNA is recognised during the S phase by *Dnmt1* (Callebaut *et* al., 1999). Homozygous mutations caused in genes coding for Dnmt1 proteins have been acknowledged to result in developmental defects (Lei *et al.*, 1996). In addition, Dnmt1 proteins have been suggested to inhibit non-CpG methylation from occurring (Ramsahoye *et al.*, 2000).

Conversely, the function of proteins in the second methyltransferases group, Dnmt2, is still unknown. *Dnmt2* has been shown to act less on DNA substrates; greater enzymatic

activity has been shown with $tRNA^{ASP}$ (Schaeffer & Lyko, 2007). Such dual activity is still unexplained. Although Dnmt2 proteins are conserved across organisms, the conservation of catalytic DNA methyltransferase motifs does not necessarily deem *Dnmt2* an active methyltransferases (Dong *et al.*, 2001), although a functional conservation has been suggested (Marhold *et al.*, 2004). For example, a *Dnmt2* enzyme has been suggested to be responsible for methylation in *Drosophila melanogaster* (Kunert *et al.*, 2003). *Dnmt3a* and *Dnmt3b* belong to the final group of methyltransferases. These two proteins have been suggested to be involved with *de novo* methylation due to their affinity for unmethylated DNA (Marhold *et al.*, 2004). *Dnmt3* enzymes have also been proposed to be involved with the establishment of methylation patterns during development (Schaefer & Lyko, 2007).

Gene regulation is achieved when epigenetic information determined by DNA methylation signals is translated into specific chromatin arrangements by methyl-DNA-binding (MDB) proteins (Schaeffer & Lyko, 2007). MDB proteins, which are the second component in methylation systems, act as a go-between between methylated DNA and chromatin modifying enzymes such as histone deacetylases, histone methyltransferases and nucleosome remodelling enzymes (Hendrich & Tweedie, 2003).

Methylation systems vary from organism to organism. For instance, comparison of the methylation system in *D. melanogaster* and vertebrates illustrates that DNA methylation in *D. melanogaster* occurs predominantly during embryonic development (Lyko *et al.*, 2000a), while methylation in most vertebrates occurs later on in life (Marhold *et al.*, 2004). Similarly, DNA methylation in *D. melanogaster* occurs in areas

of non-symmetrical CpA and CpT dinucleotides (Lyko et al., *2003)*, while DNA methylation in the latter occurs at palindromic CpG dinucleotides. Also, while DNA methylation patterns in vertebrates are established and governed by a range of methyltransferases, methylation in *D. melanogaster* appears to be determined by a single *Dnmt2* enzyme (Kunert *et al.*, 2003).

Functional homologes of Dnmt2 proteins have been determined in vertebrates such as *Homo sapiens*, *M. musculus* and *Xenopus laevis*, as well as in insects such as *Bombyx mori, Anopheles gambiae* (Clayton, 2007) and *D. melanogaster* (Dong *et al.*, 2001), all of which contain all of the 10 catalytic motifs which are characteristic of cytosine-5 methyltransferases (Marhold *et al.*, 2004).). *Dnmt2* amino acid similarities have been illustrated to be approximately 60% between drosophilids and *A. gambiae* and 50% between drosophilids and vertebrates (Marhold *et al.*, 2004). Studies that have explored Dnmt2 proteins of different dipteran species have suggested that the high level of conservation as well as the significant amount of DNA methylation within the genomes indicate that Dnmt2-based methylation has survived more than 250 million years of evolution. *Drosophila Dnmt2* genes, however, contain additional amino acids between the target recognition domain and the IX catalytic motif, which are not found in *A. gambiae* or in vertebrates (Marhold *et al.*, 2004). *Dnmt2* methyltransferases are the only such enzymes that have been found in *D. melanogaster* and *A. gambiae*, both of which exhibit methylation (Marhold *et al.*, 2004). Therefore, it has been suggested that *Dnmt2*, while highly conserved, has a catalytic function. This hypothesis has been further re-enforced by evidence of mammalian *Dnmt2* showing weak catalysis (Liu *et al.*, 2003). *Dnmt2* has been suggested to be associated with non-CpG methylation,

which makes sense considering CpA methylation has been found *A. gambiae* and *D. melanogaster*.

1.7 Methylation systems in insects

The presence of 5-methylcytosine in insects has been observed in many orders, including Diptera, Homoptera, Hymenoptera, Lepidoptera and Orthoptera (Field *et al.*, 2004) (Figure 1.1). Absences in other insect orders have also been observed, although it has been suggested that this may be due to DNA methylation being developmental stage-specific or targeting specific sequences only (Field *et al.*, 2004). However, information regarding methylation in insects is limited due to the absence of a suitable model system to conduct functional analysis. Methylation in insects contradicts the presence of a conserved function due to varying levels of methylation within different insect genomes (Field *et al.*, 2004), which may be due to high diversity within insects.

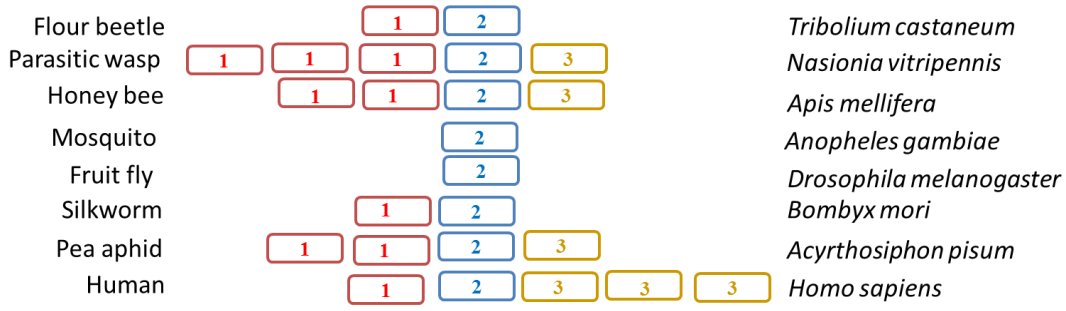


Figure 1.1: The distribution of DNA methyltransferases in humans and a range of insects. *Anopheles gambiae* and the fruit fly *Drosophila melanogaster* only have the methyltrasnferase *Dnmt2*. The flour beetle *Tribolium castaneum* and the silk worm *Bombyx mori* have *Dnmt1* and *Dnmt2*. The parasitic wasp *Nasonia vitripennis* has three orthologues of *Dnmt1* and a single orthologue of *Dnmt2* and *Dnmt3*. The honey bee *Apis mellifera* and pea aphid *Acyrthosiphon pisum* have two orthologues of *Dnmt1*and a single orthologue of *Dnmt2* and *Dnmt3*. Humans have a single *Dnmt1* and *Dnmt2* but three orthologues of *Dnmt3*.

The following section is a synopsis on studies conducted on four different insects, *Drosophila melanogaster*, *Mamestra brassicae*, *Mysuz persicae* and *Planococcus citri*, which exhibit a wide variety of effects due to DNA methylation.

DNA methylation in the fruit fly *D. melanogaster* is unique when compared with other insects. *D. melanogaster* has a low methylation level with less than 1% of cytosines being methylated, with the highest levels of methylation found in early embryos which yield inadequate quantities of DNA for biochemical and molecular manipulation (Lyko, 2001). In addition, a majority of the 5-methylcytsine occurs in non-CpG dinucleotides, therefore making conventional CpG specific techniques ineffective (Lyko, 2001). Methylation in *D. melanogaster* is instigated by an enzyme that is a member of the Dnmt2 group of methyltransferases (Kunert *et al.*, 2003). Over-expression of *Dnmt2* in *Drosophila* has been observed to increase methylation at CpT and CpA dinucleotides (Kunert *et al.*, 2003). CpA has also been found in *A. pisum* at 3 sites in an intron (Walsh *et al.*, 2010).

The presence of non-CpG methylation in *D. melanogaster* and the exclusivity of this non-CpG methylation make *D. melanogaster* very distinct from other organisms. Usually, post-replicative copying from parental to daughter strand maintains methylation patterns. A key necessity for this mechanism is the presence of symmetric CpG or CpNpG sequences (Field *et al.*, 2004). Due to lack of evidence of such symmetry being present in *D. melanogaster*, it has been suggested that an alternative mechanism retains epigenetic information, possibly in the form of MDB proteins (Hendrich & Tweedie, 2003). The MDB2/3 protein in *D. melanogaster*, which is a

homologue of vertebrate MDB proteins, has been documented to have a significant affinity towards embryonic DNA, which also corresponds to the peak of DNA methylation (Marhold *et al.*, 2002). Therefore, MBD2/3 has been suggested to sustain epigenetic information by attaching to methylated DNA and adhering to histone modifying enzymes (Field *et al.*, 2004). Knock-out experiments where Dnmt2 proteins were silenced using RNA interference in *D. melanogaster* embryos did not effect development (Kunert *et al.*, 2003), which is similar to observations in the bread mould *Neurospora crassa* and mice, where no phenotypic differences were observed (Field *et al.*, 2004).

Mamestra brassicae, the cabbage moth, has been documented to possess the highest amount of methylation in any insect genome with 10% of cytosines being methylated, which is similar to the extent of methylation found in vertebrates (Field *et al.*, 2004). Further analysis has shown that the 5-methylcytosines are not exclusive to clustered and unclustered CpG doublets, but also occur in the outer C of the 5'-CCGG-3' sequence (Mandrioli & Volpi, 2003). In addition, it was found that methylation is absent in repetitive sections of DNA (Mandrioli & Volpi, 2003). Transposons are usually significantly heavily methylated in plant and vertebrate genomes, which has led to the hypothesis that methylated cytosines is a conserved mechanism that prevents transposition (Field *et al.*, 2004).

The diversity of the functionality of DNA methylation in insects is further illustrated in insecticide resistant strains of the peach-potato aphid, *M. persicae*. This resistance to insecticides is a result of amplifying detoxifying esterase genes which results in highly

resistant individuals possessing up to eighty copies of the gene (Field *et al.*, 1999). Experiments have shown that the two alleles of this gene, E4 and FE4, have 5-methylcytosine in the CpG doublets within and around the gene (Field *et al.*, 2004). It has been suggested that the methylation was caused by the amplification of the esterase genes, similar to methylation in the presence of parasitic DNA in vertebrates and transgenes in plants. Furthermore, a simultaneous loss in E4 expression and methylation (Hick *et al.*, 1996) occurred. However, an association between methylation and loss of esterase gene expression has been found only in aphids (Ono *et al.*, 1999; Field *et al.*, 2004).

CpG profiling and bisulphite sequencing has shown that 5-methylcytosine is present in the E4 gene but is absent upstream of the gene, especially in the CpG rich region where transcription usually starts (Field, 2000), thus illustrating that methylation would not interfere with transcription. Additional findings illustrate that there is a reduction in CpG doublets within the E4 gene (Shimizu *et al.*, 1997), which demonstrates the evolutionary effect of 5-methlcytosine mutations. The maintenance of methylation over such a long period of time has been suggested to be due to methylation enhancing transcription via the prevention of incorrectly initiated transcripts by the silencing of illegitimate gene promoters (Field *et al.*, 2004).

1.8 Social insect systems & worker reproduction

Eusociality has been defined as an advanced level of social organisation, characterised by overlapping adult generations, collective brood care and the presence of a reproductive division of labour amongst members (Wilson, 1971; Crespi & Yanega,

1995). Eusociality therefore results in the generational care of another's (reproductive caste, such as queens) offspring by individuals who do not reproduce themselves (non-reproductive caste, such as workers). Worker behaviours have been suggested to have evolved from maternal care behaviours and the primitively eusocial wasp *Polistes metricus* has been used to test the theory. Results indicate that the trascriptome in worker wasps was more similar to the foundress', as opposed to queens and other gynes (Toth *et al.*, 2009).

Workers in social insect colonies are capable of laying unfertilised eggs, which develop into males (Wenseleers & Ratnieks, 2006). However, there is usually a price to pay for this behaviour. Other workers can 'police' and destroy the eggs before they hatch. Non-reproducing workers can also be aggressive towards workers with active ovaries, resulting in death or a regression in ovary development (Ratnieks *et al.*, 2006) thereby maintaining reproductive division of labour in the colony.

Worker reproduction in social insect colonies has been compared to cells which lead to tumour formation in other organisms (Beekman & Oldroyd, 2008). This form of reproduction usually comes with a high cost to the colony because reproducing workers very rarely contribute to maintaining the colony. The increase in brood can also cause problems in rearing- the number of workers may be insufficient to care for all the brood (Wenseleers *et al.*, 2004b).

A colony is not only more effective in the competition for resources, but is also invaluable for defence (Wilson, 2008) compared to solitary individuals. Altruism in social insect societies is widespread and best illustrated in how workers sacrifice their own reproduction in order to take care of the queen's offspring (Ratnieks & Wenseleers, 2007). While this seems to go against natural selection, which favours reproductive success, Hamilton's theory of inclusive fitness suggests that altruism is promoted if the two individuals in question are genetically related. However, the assumption made by Hamilton, that the individual gives up its reproductivity voluntarily is not true and sometimes altruism is coerced. Egg laying by workers, for example, should occur less in species which exhibit more efficient worker policing (Wenseleers *et al.*, 2004) if enforced altruism occurs. This would result in a reduced chance for workers to lay eggs. A study in ten species showed that this was true as instances where worker-laid eggs were killed occurred if the species had a smaller percentage of workers who laid eggs (Wenseleers & Ratnieks, 2006).

An absence of coercion would still lead to altruism and societies would still operate due to the relatedness between members. If individuals are coerced into being altruistic, it has been suggested that there should also be a strong selection for those that try to evade coercion (Ratneiks & Wenseleers, 2007). Evidence shows that Cape honeybee workers evade worker policing by chemically mimicking queen-laid eggs (Martin *et al.*, 2002). *Apis florea* and *Apis cerana* workers evade policing by laying eggs in unrelated queenless colonies (Nanork *et al.*, 2006). In addition, bumblebee workers can sometimes kill their own queen (mother) so that they can avoid queen policing and lay their own eggs (Bourke, 1994).

Coercion can have two main benefits according to inclusive fitness theory. First, it increases the colony's reproductive potential. Secondly, it increases the relatedness between the coercers and young queens and males (Ratnieks, 1988). Coercion has also been suggested to have played an important role in instigating eusociality (Ratnieks & Wenseleers, 2007). However, if this was the case, eusociality would occur more often in nature (Ratnieks & Wenseleers, 2007).

1.9 Bumblebees: a model for epigenetics

The *Bombus* genus is composed of approximately 250 eusocial or secondarily socially parasitic species. All *Bombus* colonies are annual and founded by a single queen before the first eggs are laid. Eggs tend to be laid together in clumps in an 'egg cell'. In addition, morphological differences between the castes are limited to size only (Cnaani *et al.*, 2002).

The use of bumblebees as a model species in the study of population, ecology, host-parasite interaction and behaviour is on the rise. They are a key pollinator of commercial crops, in addition to wild flowers (Goulson, 2003, Moret & Schmid-Hempel, 2000 & Riddel *et al.*, 2011). They are also used in conjunction with honeybees in comparative studies on social evolution (Mares *et al.*, 2005). The *B. terrestris* karyotype is composed of 18 chromosomes and is estimated to be 274mB in size (Baylor College of Medicine Human Genome Sequencing Center).

Bumblebees too have a reproductive division of labour, where the queen reproduces while the worker caste defends the nest, forages and helps rear the young. Males contribute very little, if anything towards the nest. Like all hymenopteran species, bumblebees are haplodiploid, in that they have a single locus sex determination system (Cook & Crozier, 1995). Queens and female workers develop from fertilised eggs, while males are a result of unfertilised eggs, produced parthenogenetically by either the queen or sometimes workers. *B. terrestris* queens, unlike honeybee queens, are singly mated (Schmid-Hempel & Schmid-Hempel, 2000) and most species have an annual life cycle.

Bombus terrestris queens emerge from hibernation in February to March and start looking for suitable nest sites. They usually always nest underground, sometimes using former rodent burrows. She then constructs a nest using materials like moss, hair and grass to form a spherical lair, with a central chamber. This nest will usually have only one entrance. The queen then moulds pollen into an egg cup and lays her eggs. Usually, 8-16 eggs are laid initially. The eggs are then covered by a mixture of wax secreted from the queen's ventral abdominal surface and pollen and the brood is incubated by the queen. After approximately four days, the eggs hatch into larvae which start to consume pollen (Goulson, 2003). In addition to incubating the brood, the queen also frequently forages in order to maintain adequate pollen and nectar stocks. The colony is at its most vulnerable point, since bad weather or insufficient foraging can cause the colony to collapse.

Bumblebee larvae have four instars. They develop for ten to fourteen days and then pupate. The pupae develop for approximately another two weeks. Although the queen incubates the larvae and pupae, those towards the centre of the brood clump receive more warmth and consequently grow more and emerge first. After the first batch of eggs develops into pupae, the queen will lay more eggs. Newly emerged workers are white and colours develop after a day. Once the first workers emerge, the queen will stop foraging and some of the workers take over the task. Others stay in the colony and help the queen with subsequent broods. The colony continues to grow and sometimes can escalate to nearly ten times its original weight in four weeks (Goulson *et al.*, 2002a). If the colony develops satisfactorily despite the dangers such as parasitism, predators and physical harm to the nest site, the queen stops laying workers and solely lays larvae destined to be new queens and males. The stage when the queen predominantly starts laying haploid eggs is called the switching point (Figure 1.2). Until the switching pointing, the workers' fecundity is stifled by pheromones produced by the queen. Workers with developed ovaries at this stage of the colony's life sometimes start to destroy the eggs laid by the queen and start to lay unfertilised eggs of their own, leading to the competition point, where there is conflict between the reproductive workers and queen due to worker egg laying. Usually, the queen 'polices' this practice and destroys non-queen laid eggs.

Young queens forage and return to the nest, continually increasing their fat reserves via the consumption of pollen and nectar. Males leave the nest. Once mated, the young queens feed and find suitable hibernation locations. At the same time, the original colony collapses. While *B. terrestris* is considered to have an annual life cycle, they also exhibit phenotypic flexibility; *B. terrestris* have been documented to maintain

colonies throughout milder winters in New Zealand, North Africa and Corsica (Goulson, 2003).

Conflicts over worker reproduction are frequent at the end of annual colonies headed by a singly mated queen. This is due to the workers being more related to other worker's sons than the sons produced by the queen (the worker's brothers). Bumblebee workers usually produce drones at the end of the colony's life after the colony has reached a sufficient size, since a greater number of workers allows for better brood care. Bumblebee colony size depends on parasite load, genetic diversity and availability of resource (Herrmann *et al*., 2007). *B. terrestris* queens control worker reproduction by pheromonally controlling egg development during the initial days after workers emerge and by preventing oviposition until the colony reaches the end of its life (Duchateau & Velthuis, 1989). Worker ovaries are initially mature and sometimes contain a few mature eggs but the mature eggs are reabsorbed. Workers usually do not reproduce until the queen has reached the switching point, which is when she starts to lay sexuals. It has been documented in *B. terrestris* that all male offspring produced by workers before the switching point were offspring of non-natal workers (Duchateau & Velthuis, 1988). Once competition point was reached, both natal and non-natal workers produced drones (Beekman & Oldroyd, 2008).

Figure 1.2:*Bombus terrestris* life cycle.

The queen emerges from hibernation, finds a suitable nesting site and starts to forage in spring. She lays her eggs, which develop into larvae and then pupae before emerging as callow workers. Callow workers are bees less than a day old. The colony develops, with the queen laying subsequent batches of workers. At the switching point, the queen starts laying new queens and males. The queen stops controlling worker reproduction simultaneously and workers with developed ovaries start to lay eggs destined to become haploid males. The destruction of worker-laid eggs by the queen leads to competition between the queen and workers. The new queens and males leave the colony to mate towards the end of summer, after which the new queens prepare for hibernation. NB: the picture for mating shows a *Bombus lapidarius* queen, while all other pictures show *B. terrestris.*

In a monoandrous species such as *Bombus terrestris*, workers are related to the own eggs (r=0.5) and are less related to their sister's eggs (r=0.375). They are related even less to their brothers produced by the queen (r=0.25). Therefore, it is expected that *B. terrestris* workers would favour their sisters reproducing over the queen. Policing at this instance is mainly conducted by the queen herself, as opposed to other workers as in the honeybee (Beekman & Oldroyd, 2008).

While the kin selection theory focuses on relatedness controlling worker reproduction, the conflict theory focuses on maternally derived and paternally derived alleles being subject to different selectional pressures. In queen-right *B. terrestris* colonies (colonies where a queen is present), the patrigene in a worker is unrelated to that in the queen's son but has a 50% chance of being present in a worker's son. This is because the queen uses unfertilised eggs to produce drones, while worker's use either the queen's genes or their father's genes to produce haploid drones. A matrigene in a worker has an equal chance of being present in either the worker's son or the queen's son (50%). As a result, genes which promote reproduction in workers in queen-right colonies would be expected to have an imprinted matrigene, while the patrigene is expressed. Conversely, in genes which supress worker reproduction in queen-right colonies, the patrigene would be imprinted and the matrigene would be expressed.

Members of the social Hymenoptera interact for their entire lives with other colony members who are closely related to themselves. Relatedness asymmetries are caused by their haplodiploid sex determination system. In addition, different species show different social systems and can lead to different conflicts within the colonies (Queller

& Strausmann, 2002). The haplo-diploid nature of bumblebees make them a suitable model organism for validating Haig's Conflict theory since each colony is composed of full siblings with different roles to play in the colony as well as the presence of different castes within the same colony. Importantly, phenotypic plasticity is exhibited by *B. terrestris*, in that the absence of a queen can lead to ovary development and the laying of haploid eggs by workers. *B. terrestris* are also hailed to be at the border of primitive and advance eusociality (Pereboom *et al.*, 2005). The colonies are also annual, which means individuals do not have a second chance at increasing their reproductive output. Thus, they are an invaluable asset in the study of differentially expressed genes in reproducing and non-reproducing individuals. Other features that promote the use of bumblebees as model organism in genetic and epigenetic studies is the ease with which they can be reared in the lab, the short generation time and the recent developments in bioinformatic resources- the *B. terrestris* BAC library and the partially annotated genome. The short generation time is advantageous if it is necessary to conduct back crosses to confirm imprinted gene expression.

1.10 Thesis aims

The main focus of this thesis was to conduct an independent test of Haig's conflict theory using *Bombus terrestris*. This test is termed independent because a social insect system is used instead of a mammalian or angiosperm system. In addition, the allocation of resources in this test is not based on the presence of a placenta or endosperm, but on the feeding of brood by non-reproducing worker bees. This means that the imprinted genes would be expected to be found in workers which exhibit phenotypic plasticity through the development of ovaries and subsequently laying eggs

when a queen is absent. The loss of the queen in a colony of bumblebees would result in a change in the worker bee's environment, in that the restraint against worker reproduction has been lifted.

Before testing Haig's theory, it is necessary to establish the presence of a means for imprinting, which in this case is a functional methylation system. The first research question which is addressed in Chapter 2 is whether bumblebee workers have a functional methylation system. Such a system would indicate a means of facilitating phenotypic plasticity and thereby genomic imprinting. Currently, the European honeybee *Apis mellifera*, which like the bumblebee is also a member of the family Apidae and a hymenopteran, the jewel wasp *Nasonia vitripennis*, pea aphid *Acyrthosiphon pisum* and flour beetle *Tribolium castaneum* have been documented to have components of mammalian-like methylation machinery (Walsh *et al*, 2010). Methylation systems across the class Insect tend to vary however. In the insect systems analysed thus far a combination of *Dnmt1*, *Dnmt2*, *Dnmt3* and a methyl-binding protein have been identified. Not all species have the same number of orthologues per methyltransferase and some seem to have lost methyltransferases. For example, *Drosophila melanogaster* lacks *Dnmt1* and *Dnmt3*.

If a methylation system is found in *B. terrestris*, the next step is to confirm that this system is functional by searching for evidence of methylated genes. Methylation profiles change with an individual's age. In addition, bumblebee colonies are composed of individuals from different castes, each with their own set of behaviours. An analysis

of methylated genes and methylation differences in *B. terrestris* caste and life stages is conducted in Chapter 3.

The final step involves testing Haig's conflict theory using worker reproduction in queen-less colonies, which is examined in Chapter 4. The main research question will be analysing if methylated genes are imprinted according to the theoretical expectations of Haig's conflict theory when applied to queen-less worker reproduction in a monoandrous social insect system. In this situation, any given patrigene in a worker has a 50% chance of being in her son and a 50% chance of being in another worker's son. However, a matrigene in a worker has only a 25% chance of being in another worker's son. Therefore, conflict theory predicts that workers' patrigenes and matrigenes are subject to different selectional pressures and, for those genes that turn on reproduction in workers, the maternal copy should be expressed and the paternal copy imprinted. If an imprinted patrigene is found in a gene associated with worker reproduction, the validity of Haig's conflict theory as an explanation of the evolution of genomic imprinting will be confirmed.

Therefore, the main aims of this thesis can be summarised as -

1. Establish the presence of a methylation system in *Bombus terrestris*.
2. Assess if this methylation system is functional.
3. Test Haig's theory using one of many conflicts found in social insect systems- worker reproduction in queen-less colonies.

Chapter 2 : Methylation machinery in *Bombus terrestris*

2.1 Introduction

Using insects as model species to explore methylation is not a novel idea. The complex nature of methylation and a high methylation percentage in mammalian genomes can impede research. Insects have relatively small genomes which are lightly methylated and unlike vertebrates, exhibit methylation almost entirely in the gene body itself; transposons and repetitive elements in insects are not methylated, unlike in mammals. (Lyko & Maleszeka, 2011).

While a more precise time scale is not known, it has been suggested that the tribes Bombini and Apini split 125-80 million years ago, concurrent with the appearance of angiosperms (Stolle *et al.*, 2011). The *Bombus* and *Apis* genera are thought to have separated considerably later. Even though the two genera independently evolved for a 100 million years, a great percentage of their genomes and even full chromosomes are highly conserved (Stolle *et al.*, 2011). The high level of conservation makes methylation studies in the honeybee *Apis mellifera* of significant importance when evaluating methylation in the bumblebee *Bombus terrestris*.

Methylation has been shown to be associated with environmentally controlled phenotypic plasticity in *A. mellifera*. Kucharski *et al.* (2008) used RNAi to silence the expression of *Dnmt3* in newly hatched L1 larvae. The L1 larval stage was used because *Dnmt3* is a *de novo* methyltransferase which establishes methylation patterns during development (Schaefer & Lyko, 2007). Embryos, another key stage of development, were not used since the RNAi treatment resulted in fatalities (Kucharski *et al*., 2008). The larvae were injected with either a control gene *uth* (untreated) siRNA or *Dnmt3*

siRNA and allowed to develop into adults. While the larvae in both groups developed like normal wild *A. mellifera* larvae, they demonstrated phenotypic plasticity at adulthood. Approximately 72% of the *Dnmt3* siRNA treated larvae emerged as queens with developed ovaries, while the remaining 28% were workers with underdeveloped ovaries of 2-6 ovarioles. Comparatively, 77% of the larvae in the control group emerged as workers while 23% of the larvae developed into morphologically queen-like individuals. These queen-like individuals had underdeveloped ovaries with 50-80 ovarioles per ovary, compared to the 120-190 ovarioles found per ovary in queens that developed from *Dnmt3* siRNA treated larvae, illustrating that DNA methylation plays a major role in the regulation of queens in *A. mellifera*. The ovaries of siRNA induced queens were identical to those found in virgin queens reared on royal jelly. Phenyl butane, a component of royal jelly, is a histone deacetylase (HDAC) inhibitor. Histone deacetylaces lead to the compacting of chromatin, which causes transcription repression (Chittka & Chittka, 2010). Therefore, the silencing of *Dnmt3* and diet of royal jelly cause a phenotypic change which promotes ovariole development.

Methylation has also been suggested to play a part in social interactions, long term memory (Lockett *et al.*, 2010) and communication in *A. mellifera* (Gabor *et al.*, 2010). Bees are the only other species excluding primates known to have evolved a symbolic language. The high cognitive capacity needed for the development, neurology and behaviour of bees may be influenced by epigenetics (Lyko & Maleszeka, 2011) by assisting the integration of innate signals and environmental signals with the use of DNA methyltransferases, histone deacetylases and methyl-binding proteins (Goll *et al*., 2005).

Environmental changes during the development of larvae result in contrasting anatomical and behavioural characteristics in honeybee castes. In addition, the study by Kucharski *et al.* (2008) discussed above illustrates that methylation in *A. mellifera* is an important component of an epigenetic network which controls reproductive division of labour. Reproductive division of labour is a key feature of eusociality (Drapeau *et al*., 2006). Honeybees are the closest phylogenetic species to bumblebees shown to possess a full complement of methyltransferases. If worker reproduction in *B. terrestris* is epigenetically controlled, it must first be known if a system which facilitates epigenetics exists in *B. terrestris.*

Prediction:

The bumblebee *Bombus terrestris* like the honeybee *Apis mellifera*, has full complement of DNA methyltransferases.

2.2 Methods

The aim of this chapter is to establish the presence of a methylation system in *Bombus terrestris*. Two methods were used to determine the presence of methyltransferases in *B. terrestris*. First, *B. terrestris* DNA was amplified using primers spanning conserved sections of other insect methyltransferases (2.2.1). Then the *B. terrestris* BAC library was screened (section 3.2.2). Finally, an analysis of the structure, sequence conservation and evolution of *B. terrestris* methyltransferases was conducted using bioinformatics (section 3.2.3).

2.2.1 Bee keeping

Bumblebee colonies were purchased from a commercial supplier (Koppert; Haverhill, U.K.) and reared in wooden nest boxes (inner dimensions- 24 x 16 x 13.5cm) (Fig. 2.1) under constant darkness with red light conditions at 26°C and 60% humidity. Bees were fed *ad libitum* with pollen and 50% diluted glucose/fructose apiary solution (Meliose; Roquette, France), which contained Nipagine fungicide (16.2g of Nipagine dissolved in 135ml of absolute ethanol). Apiary syrup was provided via a perforated centrifuge tube (Anachem Limited) in a separate smaller clear Perspex box (17cm x 5.5cm x 11.5cm) with cat litter, which was connected to the main nest with a plastic tube (Riddell *et al*., 2009).

Figure 2.1: Nest set up. The bee colony was reared in the wooden nest box with pollen provided in a petri dish. Apiary solution was provided in the connected Perspex box.

2.2.2 DNA extraction

5% (w/v) Chelex was prepared by heating 2.5g of Chelex powder with 50ml of distilled water using a magnetic stirrer on a hotplate stirrer until the bottle was too hot to touch (approx. greater than 50^{o}C). A single leg removed from five adult worker bees of the same colony was crushed using liquid nitrogen with a pestle and mortar and placed in an Eppendorf tube. 300µl of Chelex was added and the sample was incubated at 90°C in a thermoblock for 10 minutes. The sample was vortexed every two minutes during the incubation period and placed on ice to cool. The sample was then centrifuged for 1 minute at maximum speed of 13,000rpm (Progen GenFuge 24D bench-top centrifuge). If the sample was being used afters torage in a freezer, it was defrosted and centrifuged for 5 minutes.

Genomic DNA was also extracted from 10mg of leg tissue from three adult workers of the same colony using the Promega Wizard® SV Genomic Purification System. Instructions were followed according to the manual.

2.2.3 DNA quantification and storage

The equation below was used to calculate DNA concentration using a mass spectrophotometer at 260nm. All DNA samples were stored at -20°C.

$$\text{Concentration (µg/ml)} = A_{260} \text{ reading} \times \text{dilution factor} \times 50\text{µg/ml}$$

2.2.4 Construction of primers for PCR

Primers to amplify *B. terrestris* DNA were designed using the *A. mellifera* methylation gene sequences (Table 2.1) with primer construction programs Invitrogen's Oligoperfect™ Designer (http://tools.invitrogen.com/content.cfm?pageid=9716) and Primer 3, version 0.4.0 (http://frodo.wi.mit.edu/).

Table 2.1: *A. mellifera* methyltransferase gene information. The accession numbers and size for methyltransferase genes and the methyl-binding domain gene in *A. mellifera*

GENE	ACCESSION NUMBER	LENGTH
AmDnmt1A	GB15130	5691bp
AmDnmt1B	GB19865	4137bp
AmDnmt2	GB10767	1487bp
AmDnmt3	GB14232	5116bp
AmMBD	REFSEQ XP_392422	2017bp

Sections upstream and downstream of the genes were also included so that the entire gene could be amplified. Due to the length of the genes, several primer pairs were designed to overlap sections of the genes. Primers were ordered from Sigma Aldrich and stock solutions were made to 100pmol/µl; working solutions of 10pmol/µl were used. Primers were stored at -20˚C.

2.2.5 Amplification & visualisation of DNA

Optimised PCR reaction volumes of 25µl were composed of 12.5µl of YB-Taq 2xbuffer (YB), 2µl of each primer, 2.5µl of DNA, 2.5µl of 10mM $MgCl_2$ and 3.5µl of water. PCR reactions were then run on either a 96 well T1 thermocycler (Biometra), a T Professional Gradient thermocycler (Biometra) or a MultiGene 2 (Labnet). Lids for all PCRs were maintained at 100°C and a hold step at 4°C followed the PCR programme.

The PCR cycle consisted of 30 3-step cycles, with 30s at 94°C, 1.5 minutes at primer-specific annealing temperatures and 1.5 minutes at an extension temperature ranging from 70-74°C. A final extension period of 10 minutes followed.

Agarose gels for electrophoresis were made by dissolving 1% - 3% agarose (Melford, UK) with TAE (EDTA, acetic acid and ddH_2O, prepared according to Sambrook *et al.*, 1989) in a microwave, stopping occasionally to swirl the contents until the agarose had melted. The agarose was cooled under running water before 1% ethidium bromide (Fisher Scientific, UK) was mixed in and then poured into gel trays with combs and were left to set for 30 minutes. The gels were run in 1x TAE at 80- 120V using a Biometra powerpack (Standard Powerpack, P25) for 30-60 minutes along with DNA ladders (Q Step 1, YB, UK; Hyperladder 1, Bioline, UK, 50bp or 100bp ladders N.E.B). Gels were visualised on a GeneFlash (Syngene, UK) transilluminator system. 1µl of loading buffer was added per 5µl of sample that was loaded. The DNA bands were visualised under UV light on a transillumiator.

2.2.6 Purification and sequencing of PCR products

All reactions were repeated as a 50µl reaction volume and the products were then excised from the gel after electrophoresis. The Promega Wizard®SV Gel and PCR Clean-Up System was used to purify the gel bands. The PCR products were sent to the John Innes Centre (Norwich) for type 2 commercial sequencing.

2.2.7 The *Bombus terrestris* BAC library

This experiment was carried out by Gillian Dornan, a MSc student, under my lab supervision.

A BAC library for *B. terrestris* was ordered from Clemson University Genomics Institute (South Carolina, USA). 36,864 BAC clones were distributed through 4x96 plates, producing a total of 384 wells. The vector pIndigoBac536 was used, which has a chloramphenicol antibiotic marker. The average insert size was 120 kb. Pre-dotted filters, Filter A and Filter B, were also ordered, on which the BAC DNA was arranged.

2.2.7.1 Synthesis and radiolabelling of probes

Probes were created using primers which amplified between 200-400bp sections of DNA using results from section 3.2.1. The probes were amplified, electrophoresed on a 1% gel, excised from the gel and purified using the Wizard SV Gel and PCR Clean Up kit (Promega, UK). The sample was then sent to the John Innes Sequencing Centre to confirm that the fragment was from a specific methyltransferase.

The DECAprime™ II Kit (Ambion, USA) was used with [α-32P] dCTP to radiolabel the probes. 25-100ng of probe was combined with 0.37MBeq of [α-32P] dCTP according to the instructions.

2.2.7.2 Screening the BAC library using Southern hybridisation

Southern blotting allows a specific fragment of DNA on a pre-dotted filter to be detected using a radio-labelled probe which exclusively targets the fragment of DNA in question (Brown, 2001). The results can be visualised on X-ray film.

Purified probe DNA was fixed onto the pre-dotted filters using a UV transilluminator. The recipe for hybriding solution from Sambrooke (2003) was altered- 6x SSC, 1.5% SDS, 5x Denhardts solution, 0.1 M EDTA and a higher SDS concentration (above 1.2%). In addition, Dextran sulphate was not used.

Initially, the BAC filters were pre-hybridised overnight at 65˚C; subsequent pre-hybridisations were carried out for only 3-4 hours. The radioactive probe was added and the hybridised again overnight for 16-18 hours at 65˚C. The filters were then rinsed in wash buffer (1x SSC and 0.1% SDS) at room temperature and then incubated in the wash buffer for an hour at 65˚C. The filters were then sealed in a plastic bag, placed in autoradiography cassettes with x-ray film (Fuji medical x-ray film, UK) at room temperature or at -80˚C. Exposure times ranged from overnight to several days.

3 ml of LB broth was inoculated and incubated for 6 – 8 hours in a 37˚C shaker at 200 rpm. Overnight cultures for minipreps were then made by using this to inoculate 50ml of LB in a cylindrical flask. The overnight cultures were incubated again for 6 – 8 hours at 37˚C in a shaker. Cell mass was checked by calculating OD600 values. If an appropriate value was achieved, approximately 40ml was centrifuged for 10 minutes to pellet the cells. The supernatant was poured off and the BAC DNA was isolated using

either the BACPhase Prep microprep or miniprep kit (Sigma, UK). The purified BAC DNA was dissolved in ddH_2O overnight at 4˚C.

2.2.8 Analysis of methyltransferase genes

Sequences were analysed using CLC Sequence Viewer (version 6.6.1), Finch TV (Version 1.4) and Geneious Pro (Version 5.5.6). Conserved domain searches were carried out using the InterProScan (Version 36.0) plug-in on Geneious and interpreted using the Conserved Domain search on NCBI.

2.3 Results

2.3.1 PCR amplification of *Bombus terrestris* methyltransferases

The amplification of the *B. terrestris* methyltransferase genes with primers made using *Apis mellifera* gene data was successful, with sections of each gene being sequenced, except for *Dnmt1a*. Tables 2.2 to 2.7 detail the temperatures that were used with each primer pair to obtain a product of the expected size. The conditions used to amplify DNA had to be adjusted if the product was not the size it was expected to be. For example, if the PCR product obtained was shorter than the target, the annealing time was increased. If the product was longer than expected, the annealing time was shortened. If the amplification was weak, as evident by faint bands, the amount of primer used and the number of cycles was increased. However, not all primer pairs worked although a range of annealing and extension temperatures, different primer and $MgCl_2$ concentrations and cycle numbers were used.

Table 2.2: Primers used to amplify *Dnmt1a*. '-' in the annealing temperature (T^A) and extension temperature (T^E) sections are to denote primers that did not work.

PRIMER	SEQUENCE	PRODUCT	T^A/°C	T^E/°C
1ADnmt1A-F	GCGGTACAACCGTCAGCATA	1214 bp	-	-
1ADnmt1A-R	GCACACCTCCTTCTGGACAT			
1BDnmt1A-F	TGTTCTTTTGACACTGGTTTGA	1853 bp	-	-
1BDnmt1A-R	TTCTTCTCTTCTTCCCGATCA			
1CDnmt1A-F	GTCATCGGTTTTGTCCTGCT	1698 bp	-	-
1CDnmt1A-R	GTTGACATGCTCGCTTGCTA			
1DDnmt1A-F	GCAAGCGAGCATGTCAACTA	1558 bp	-	-
1DDnmt1A-F	TTGCATAACAATATGCATCACA			

None of the primers made using *A. mellifera Dnmt1a* (Table 2.2) amplified *B. terrestris* DNA. However, primer pairs D1b.1.2 and D1b.1.4 (Table 2.3), which were made to amplify *Dnmt1b* amplified a section of *Dnmt1a* as well as a section of its original target, *Dnmt1b.*

Table 2.3: Primers used to amplify *Dnmt1b*. '-' in the annealing temperature (T^A) and extension temperature (T^E) sections are to denote primers that did not work.

PRIMER	SEQUENCE	PRODUCT	T^A/°C	T^E/°C
D1b.1.1-F	TCCAGGTCATCCTAACAATGC	612 bp	-	-
D1b.1.1-R	GGATCTTCTGGTCCTGCTGA			
D1b.1.2-F	TCAACAGCCAGATTGTGGAG	733 bp	59	70
D1b.1.2-R	AGCAGGACAAAACCGATGAC			
D1b.1.3-F	GATCCTGAATGCCCAAGAAA	673 bp	-	-
D1b.1.3-R	CCCTTTCCTTTTCCACTTCC			
D1b.1.4-F	TGCATGTGCATATCGACTCA	804 bp	55	77
D1b.1.4-R	TTCGGTAAATCTCGCCAATC			

Dnmt2 is the most conserved of all the methyltransferase genes (Wang *et al.*, 2006), leading to the entire gene being sequenced (Table 2.4)

Table 2.4: Primers used to amplify *Dnmt2*. '-' in the annealing temperature (T^A) and extension temperature (T^E) sections are to denote primers that did not work. All the primers were used to amplify *B. terrestris* DNA.

PRIMER	SEQUENCE	PRODUCT	T^A/°C	T^E/°C
AmDnmt2A-F	TGTGGATGTAAATAAACAGATAGAA	1324 bp	-	-
AmDnmt2A-R	CGTTAATTGAATTTCCCAATAATC			
AmDnmt2B-F	CAGTGGTATTGGTGGAATGC	610 bp	55	70
AmDnmt2B-R	GGAATGCCAAATTGACAAGG			
AmDmnt2C-F	TGTTCTTTCATTGATTCCTCAAA	731 bp	-	-
2ADnmt2-F	AAACGAAGCCTACCCTCCAT	1580 bp	-	-
2ADnmt2-R	TTGTGATCCTTCAACATGCTCT			
2BDnmt2-F*	CCCTTGTCAATTTGGCATTC	1149 bp	55	70
2BDnmt2-R	TCGTGATATGGAAGCTATGCA			
BtDnmt2.1-F	CATCCTCGTTCCTTCATGTTC	626 bp	55	70
BtDnmt2.1-R	CAAACAACCAAGCTCTCTTTTG			
AmDnmt2.2-F	ACTGCATTTGGGTGGTGTCT	904 bp	55	77
AmDnmt2.2-R	CAACTGCTGCAACAATATCACC			
BtDnmt2.2-R	CAAGTGCCGCAACAACATCTCC			
Dnmt2.3-F	CGTTAAAAAGATTGTCGAAATAGC	748 bp	57	74
Dnmt2.3-R	CTGTTCCCCGTTTTGAAGAT			

Table 2.5: Primers used to amplify *Dnmt3*. '-' in the annealing temperature (T^A) and extension temperature (T^E) sections are to denote primers that did not work.

PRIMER	SEQUENCE	PRODUCT	T^A/°C	T^E/°C
D3-1.1-F	CAACGCACGACTTTCTTTCA	1045 bp	-	-
D3-1.1-R	GACGGGAGCAATAATCCTGA			
D3-1.2-F	CCCGACGAAAGACATCTAGG	756 bp	-	-
D3-1.2-R	CGTGTACGATACGACGCAGA			
D3-1.3-F	TGCGCGACTACACGAGCAACA	701 bp	-	-
D3-1.3-R	ACGAGATGGGACGAGGCGAA			
D3.1.4-F	AGAGGTACGAACAAAGCATCTGTCCT	665 bp	-	-
D3.1.4-R	TCCAGGGTTTACTGCACCGCCT			
D3.1.5-F	TGCTTTTCGGTTTGCTCTGTGTCT	453 bp	58	72
D3.1.5-R	TTGCACGGCGATCCACCGAT			
D3.1.6-F	TCGGCCTCCACGGTTCGAATA	538 bp	-	-
D3.1.6-R	TCTGTGCTGGGGCGAGAAGT			
D3.1.7-F	GCTACGCCCTTTATGATTCG	733 bp	53	74
D3.1.7-R	CCATTTCGTCGTTGGAAACT			
D3.1.8-F	ATGCTTTCAGGTTTGGTCGT	519 bp	-	-
D3.1.8-R	GGTAAACCATCAACGCGTCTT			
D3.1.9-F	TTCTGAAACTGGGCTTCATCT	757 bp	-	-
D3.1.9-R	GAGCAAACAGGAAAGCATTTG			
D3.10-F	ATGCCCAGCGAGTACAGATT	764 bp	-	-
D3.10-R	ATCCCTCGATTTCCCACGATA			
D3.11-F	GAGGACAGATGCTTTGTTCG	1104 bp	-	-
D3.11-R	AAGCAAACCTGAAAGCATCA			
D3.12-F	GCTCGTACTTCTGAAACTGGG	765 bp	-	-
D3.12-R	GAGCAAACAGGAAAGCATTTG			
D3.13-F	TGCAACAAAGATTGGACCAT	753 bp	-	-
D3.13-R	AAAAATTGGACCGAATGACG			
D3.14-F	TCAGGATTATTGCTCCCGTC	413 bp	-	-
D3.14-R	GCACAAAGATCCCTCGAAGA			
D3.15-F	TCCTCTCGATCGACCAGTTT	816 bp	58	72
D3.15-R	GCCCAATTTCAGAGGTACGA			

While a few primers for *Dnmt3* worked (Table 2.5), the sequences they produced were of poor quality or appeared to have amplified random sections of DNA.

Table 2.6: Primers used to amplify *MBDP*. '-' in the annealing temperature (T^A) and extension temperature (T^E) sections are to denote primers that did not work.

PRIMER	SEQUENCE	PRODUCT	T^A/ ˚C	T^E/ ˚C
AmMBP1.1-F	TCATGTTCCCGGTCGTTAAT	651 bp	-	-
AmMBP1.1-R	GGTTGAAGAATGTCGGGAAA			
AmMBP1.2-F	TTGGGAAAAACGTTTGGAAG	490 bp	55	74
AmMBP1.2-R	CTATCGAGACGGCCTGAAAC			
AmMBP1.3-F	TTCGATGTGCTTGTGGAATC	638 bp	56	74
AmMBP1.3-R	CGTGACTGGTTGAGACGAAA			
AmMBP1.4-F	TCGCGGAGTCAGAAATGAT	720 bp	56	74
AmMBP1.4-R	TATGAAACATTCACAGATAACAAAAA			
Mbp.5-F	TCATGTTCCCGGTCGTTAAT	651 bp	-	-
MBP.5-R	GGTTGAAGAATGTCGGGAAA			
MBP.6-F	ATATCCCTTTAAACATAAATGTAATGC	604 bp	-	-
MBP.6-R	TCTTATATAGACAGCGAAATAATGAA			
MBP.7-F	ACAAACGGGTTCGAAGACAG	322 bp	54	72
MBP.7-R	CTATCGAGACGGCCTGAAAC			
MBP.8-F	GCTGGCTCATGTTCCCGGTCG	429 bp	-	-
MBP.8-R	TGCATCATTTCTGACTCCGCGACT			
MBP.9-F	GTCGCGGAGTCAGAAATGAT	580 bp	-	-
MBP.9-R	GATTCCACAAGCACATCGAA			
MBP.10-F	TGCCCCAAGATCTTTAACTG	746 bp	-	-
MBP.10-R	TATATGGACCAACGGGCTTC			
MBP.11-F	TCATGTTCCCGGTCGTTAT	890 bp	-	-
MBP.11-R	CGAATGATTTGATGATTGAACC			
MBP.12-F	AATGCCCCAAGATCTTTAACTG	723 bp	-	-
MBP.12-R	TTTTGGTAAGTCCATCGCGT			
MBP.13-F	CAGCGAATATATTTTGCTTTTG	257 bp	-	-
MBP.13-R	CAACGGGCTTCAAACTTTTT			
MBP.14-F	ATTGTTCGATGTGCTTGTGG	612	55	70
MBP.14-R	AACGCAGTTGCAACAACACTCTG			
MBP.15-F	TTTCCCGACATTCTTAACC	942	-	-

The *B. terrestris* MBP protein gene was another that was almost fully sequenced (Table 2.6). A gradient PCR was carried out for primers MBP.5, MBP.6, MBP.7, D3.8, D3.9 and D2.3. To conduct a gradient PCR, each lane in the thermocycler is programmed to a different annealing temperature. The optimum annealing temperature for a primer set can be identified when identical PCR samples are placed on a number of lanes as each

sample is incubated at a different temperature during the same programme. Table 2.7 shows the range of annealing temperatures used. The extension temperature was kept constant at 72 ˚C. However none of the primers yielded results, possibly due to the locus not being conserved in *B. terrestris*.

Table 2.7: Conditions for Gradient PCR 1. The annealing temperatures used are shown in the first column. 'X' denotes the temperatures selected for each primer pair.

T^A/ ˚C	MBP.5	MBP.6	MBP.7	D3.8	D3.9	D2.3
62.0						
61.8						
61.2	X		X	X	X	
60.4	X		X	X	X	
59.5				X		
58.5	X		X	X	X	X
57.5	X	X	X	X	X	
56.5	X	X	X	X	X	X
55.6	X	X	X		X	
54.8	X	X	X		X	X
54.2		X				X
54.0		X				X

A second gradient PCR was conducted for primers MBP.11, MBP.12, D3.10, D3.11, D3.12, D3.13 and D3.14. The extension was kept constant at 72˚C, while the annealing temperatures ranged from 60±2 ˚C. Again, none of the primers yielded results.

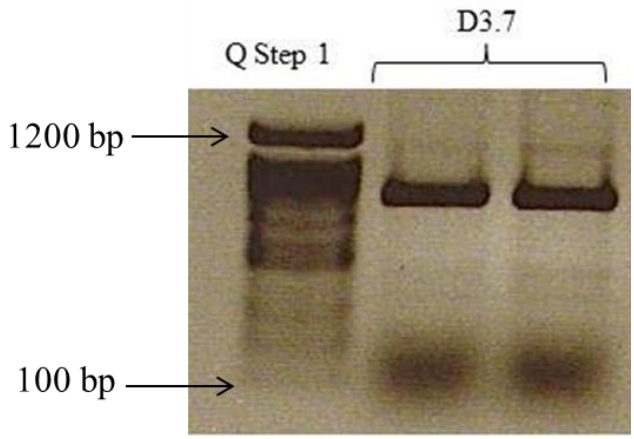


Figure 2.2: Results for primer pair D3.7. The expected product size was 733bp, which is approximately the size of the bands.

Figure 2.2 shows the results for primer set D3.7, which was used to amplify a section of *Dnmt3*. These bands were then cut out, and cleaned. Figure 2.3 shows another set of bands from PCRs used to amplify sections of *Dnmt2*, *Dnmt1b* and *MBDP*. The initial bands were cut out and cleaned and then run on a gel to confirm the presence of products, as evident by the presence of bands in the gel. These were then sent to be sequenced. Only two gels pictures (Figure 2.2 & 2.3) with products of methyltransferase genes were included in the thesis.

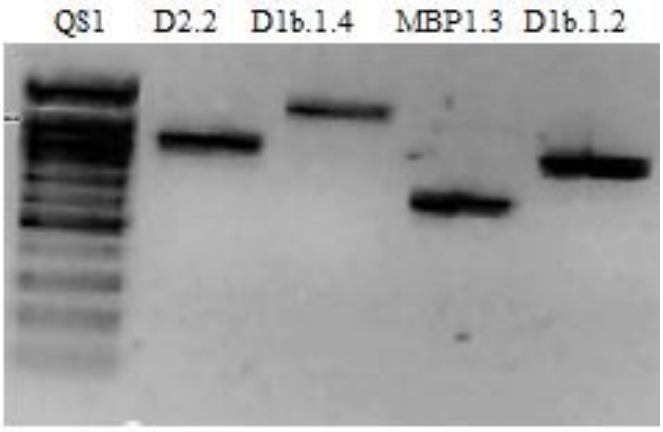


Figure 2.3: Confirming amplification. The gel shows the results of purified PCR products amplified with primers Dnmt2.2-F & R, Dnmt1b.4-F & R, MBP.3-F & R, Dnmt1b.2-F & R

2.3.2 BAC library probing

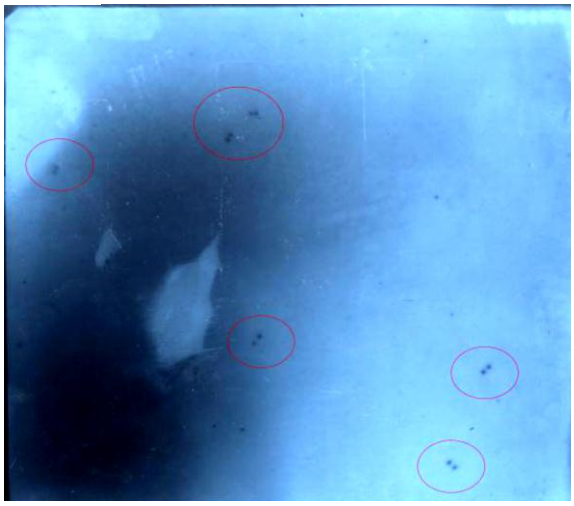
Table 2.8 shows the primers used to screen the BAC library for each gene.

Table 2.8: Primers used to screen the BAC library. Each probe was less than 300 bp in length and composed of a primer pair made using *B. terrestris* sequences from section 2.3.1

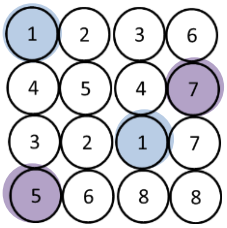
GENE	PRIMER	LENGTH (bp)
Dnmt1	F-TCAACAGCCAGATTGTGGAG	200
	R-TGATCTGAATCATCTGCTTCCT	
Dnmt2	F-GCGGCACTTGATATAAATACTGT	273
	R-CATTTCTCTTTCAGAATTTTCAA	
Dnmt3	F-TCAGGATTATTGCTCCCGTC	260
	R-CTCACAGCGCGACATTGTAT	
MBDP	F-ACAAACGGGTTCGAAGACAG	295
	R-CTATCGAGACGGCCTGAAAC	

Figure 2.4 A shows Filter B when probed for the *Dnmt3* gene. A hybridisation does not always denote that the probe bound to it's target gene. Positive hybridisations, or 'hits' are circled in red. Figure 2.4 B is an example of the clone pattern arranged on the filter. In order for a hit to be classified as 'positive', it had to have a corresponding second hit within the square. Figure 2.4.C shows the number of positive hits per gene.

A



B



C

GENE	POSITIVE HITS
Dnmt1	31
Dnmt2	14
Dnmt3	12
MBDP	11

Figure 2.4: Identifying hybridisations. A is an X-ray film of Filter B, where pairs of hybridised probes used to screen for *Dnmt3* are circled. B is an example replication pattern. Non-specific hybridisation is observed when a pair of clones which do not correspond are seen on the X-ray film (purple). When a 'positive hit' occurs (blue), the corresponding clone has also been hybridised by the probe. C shows the number of positive hits per gene.

The *B. terrestris* BAC library consists of 36,864 clones and has a high coverage, with a 99.7% chance of a sequence being successfully probed from at least one clone (Wilfert *et al.*, 2009). However, while a number of positive hits were recorded they could not be sequenced due to sequencing issues, such as degradation of the sequence.

2.3.3 Sequences, alignments, trees- a comparative analysis

The preliminary *B. terrestris* genome was published on NCBI in July 2011 and an unannotated version of the *B. impatiens* genome soon followed. The gene loci produced

by PCR and screening of the BAC library were completed with the sequences on NCBI and compared with other insect methyltransferases. Figure 2.5 is an amended version of Figure 1.1, where the *Bombus terrestris and Bombus impatiens* methylation orthologues have been added. Again, *Dnmt2* is the common denominator of methyltransferases as it is found throughout all the species.

Figure 2.5: Distribution of methyltransferases. The distribution of methyltransferase genes and the evolutionary relationships between some insect genomes, as well as the human genome is shown. Different DNA methyltransferase genes are shown using boxes of different colours. The number of boxes per colour reflects the number of orthologues in each species. The phylogeny tree was constructed using Interactive Tree of Life (http://itol.embl.de/other_trees.shtml).

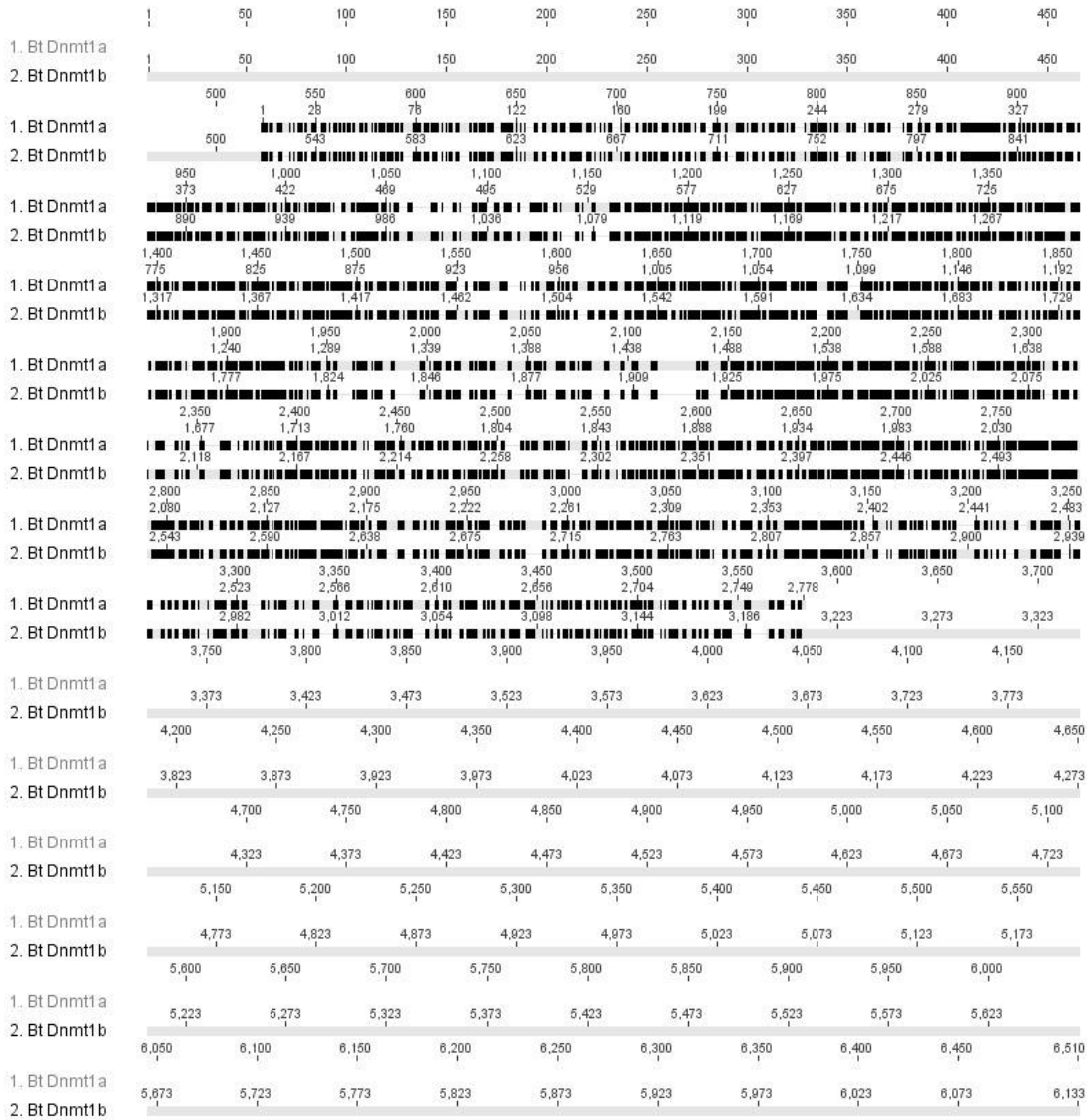


Figure 2.6:Sequence alignment for *Dnmt1a* and *Dnmt1b* in *B. terrestris*.

Figure 2.6 shows a DNA alignment of the two *Dnmt1* orthologues found in *B. terrestris, BtDnamt1a* and *BtDnmt1b*. While *BtDnmt1b* is twice as long as *BtDnmt1a*, a considerable section of both genes are similar (amino acids- 392/1447; identitiy-27%). This sequence similarity explains why primers designed for *Dnmt1b* amplified sections of *Dnmt1a*.

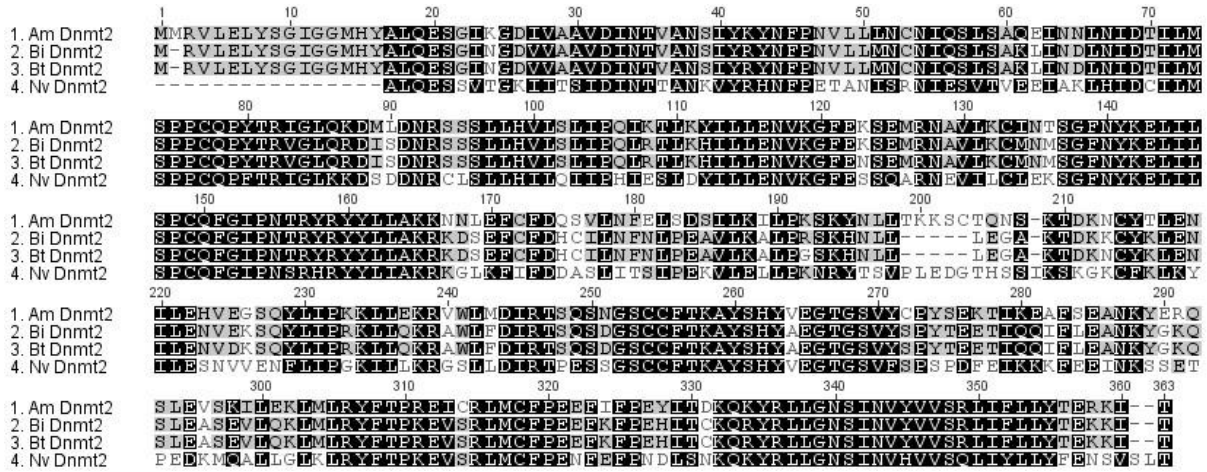


Figure 2.7: Conservation of *Dnmt2*. Protein alignment between *Apis mellifera* (*AmDnmt2*), *Bombus impatiens* (*BiDnmt2*), *Bombus terrestris* (*BtDnmt2*) and *Nasonia vitripennis* (*NvDnmt2*).

Figure 2.7 is a protein alignment of *Dnmt2* between *A. mellifera*, *B. terrestris*, *B. impatiens* and *N. vitripennis*, constructed with sequences obtained from NCBI. The gene is quite conserved across all four species. There is a 98.6% similarity between the two *Bombus* versions of the gene. Both these genes share 80.1% similarity with the *AmDnmt2*. *BtDnmt2*, *BiDnmt2* and *AmDnmt2* are approximately 50% similar to *NvDnmt2*.

Next, neighbour joining analyses were conducted using Geneious Pro (version 5.5.6). Neighbour joining analysis creates maps which illustrate genetic distance (Saitou & Nei, 1987). The gene IDs for the protein sequences used to conduct the maps were obtained from NCBI and are summarised in table 2.9.

Table 2.9: Gene accession numbers of the proteins used to conduct neighbour joining analysis. *Apis mellifera* (*AmDnmtX*), *Bombus terrestris* (*BtDnmtX*) *Bombus impatiens* (*BiDnmtX*), and *Nasonia vitripennis* (*NvDnmtX*) protein sequences were used.

GENE	GENE ID	GENE	GENE ID
AmDnmt1a	NP_001164522.1	*BiDnmt1b*	XP_003489082.1
AmDnmt1b	XP_00112269.1	*BiDnmt2*	XP_003485152.1
AmDnmt2	XP_3933991.3	*BiDnmt3*	XP_003490117.1
AmDnmt3	NP_001177350.1	*BiMBDP*	XP_003490379.1
AmMBDP	XP_392422.2	*NvDnmt1a*	NP_00116452.1
BtDnmt1a	XP_003402324.1	*NvDnmt1b*	XP_0016600175.2
BtDnmt1b	XP_003398214.1	*NvDnmt1c*	XP_001607336.1
BtDnmt2	XP_003400893.1	*NvDnmt2*	XP_001602026.2
BtDnmt3	XP_003393693.1	*NvDnmt3*	XP_001599223.2
BtMBDP	XP_003396186.1	*NvMBDPa*	NP_001164526.1
BiDnmt1a	XP_003493144.1	*NvMBDPb*	NP_001164527

Figure 2.8 shows the conservation of methyltransferases in *A. mellifera* (AmDnmtx), *B. impatiens* (BiDnmtx), *B. terrestris* (BtDnmtx) and *N. vitripennis* (NvDnmtx) are. The *Dnmt1a* genes are closely related amongst the Apidae members, as are the *Dnmt1b* and *Dnmt2* genes. While the distance between Apidae and *N. vitripennis* is expected, there is one additional finding. Figure 2.5 illustrated that *N. vitripennis* has 3 isoforms for the gene *Dnmt1*. Figure 2.8 shows that *NvDnmt1c* evolved separately. In addition, it is evolutionarily closer to the *Dnmt2* and *Dnmt3* genes than the *A. mellifera, B. terrestris* and *B. impatiens Dnmt1* genes. *NvDnmt3* diverged away from the other *Dnmt3*s early on.

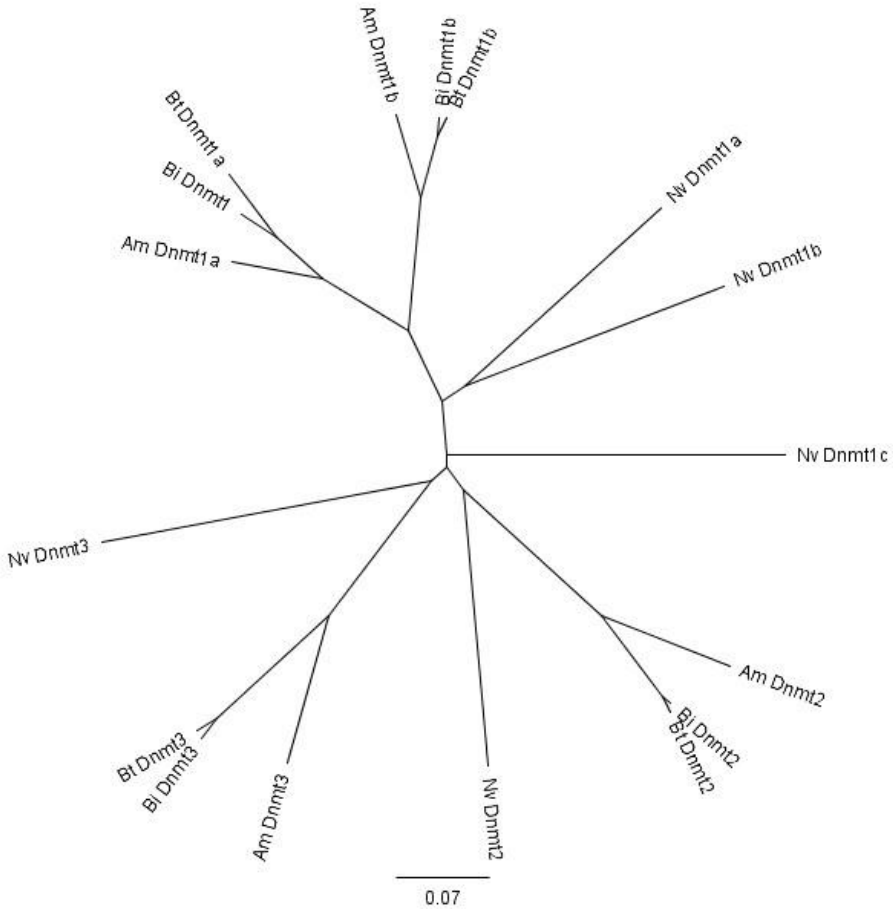


Figure 2.8: Neighbour joining analysis of methyltransferase genes between *A. mellifera* (*AmDnmtX*), *B. impatiens* (*BiDnmtX*), *B. terrestris* (*BtDnmtX*) and *N. vitripennis* (*NvDnmtX*).

The second neighbour joining tree illustrates the relationship between the *MBDP* in the same species (Figure 2.9). There are two separate clusters. As expected, the *Apidae* are clustered together, with the *Bombus* genes being closely related. The long branch leading to *NvMBDP* illustrates that it underwent a period of rapid sequence evolution.

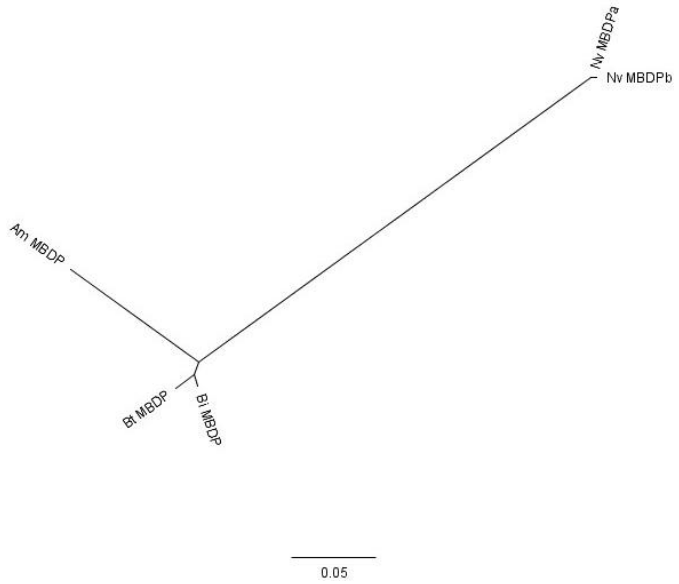


Figure 2.9: Neighbour joining analysis of MBDP. The evolutionary distance between *A. mellifera* (*AmMBDP*), *B. impatiens* (*BiMBDP*), *B. terrestris* (*BtMBDP*) and *N. vitripennis* (*NvMBDP*) is illustrated.

Finally, the conserved domains and secondary structure of the *B. terrestris* proteins was analysed using Geneious.

A number of conserved domains for *Dnmt1a*, *Dnmt1b*, *Dnmt2*, *Dnmt3* and *MBD* were found in *B. terrestris.* InterProScan used a number of databases- PANTHER, Pfam (Protein families), PIR (Protein Information Resource), PRINTS, PROFILE, PROSITE, SMART (Simple Modular Architecture Research Tool), Superfamily and TIGRFAM

(The Institute for Genomic Research's database of protein families)- to generate the data.

Figure 2.11 illustrates the domains found in *Dnmt1a*. All these domains were also found in *Dnmt1b*. The cytosine-specific methyltransferase domain (PTHR10629) found in both *Dnmt1a* and *Dnmt1b* is involved in methyltransferase activity – it catalyses the transfer of a methyl group. It also interacts with nucleic acids. The DNA (cytosine-5)-methyltransferase 1 (dnmt1) domain (PTHR10629:SF10) has a similar function (Panther Classification System). The cytosine specific DNA methyltransferase replication foci domain (PF12047) targets DMAP1 and HDAC2 during the S phase of mitosis. It assists *Dnmt1* to methylate only cytosine bases (Rountree *et* al., 2000). The BAH domains (PF01426, SM00439 and PS51038) have been suggested to connect the processes of DNA methylation, replication and transcriptional regulation (Callebaut, 1999). The CXXC zinc finger domains (PF02008 and PS51058), composed of eight conserved cytosine residues bound to two zinc ions, bind to non-methylated CpG dinucleotides (Bestor, 1992).

Figure 2.12 illustrates the conserved domains found in *Dnmt1b*. *Dnmt1b* shares all of the *Dnmt1a* conserved domains (SM00439, PS51058, PS51038, PTHR10629, PTHR10629:SF10, PF01426, PF02008 and PF12047) but also have a few in common with *Dnmt2*. TIGR00675 is another DNA-cytosine methyltransferase domain. *Dnmt1b* also contains a C-5 cytosine-specific DNA methylases C- terminal signature (PS00095) as well as a cytosine-specific DNA methyltransferase signature (PR00105). This domain provides a marker for methyltransferases which only act on cytosine (Kumar *et al.,* 1994). The S-adenosyl-L-methionine-dependent methyltransferases domain

(SSF53335) is involved in the formation of 5-methylaminomethyl-2-thiouridine by catalysing the transfer of methyl groups from S-adenosyl-L-methionine to 5-aminomethyl-2-thiouridine. The C-5 cytosine-specific DNA methylase (PF00145) also catalyses the transfer of a methyl group. C5_MTASE_1 **(**PS00094) is a C-5 cytosine-specific DNA methylase active site. DNA (cytosine-5)-methyltransferase 1, DNMT1 type (PIRSF037404) is specific to only *Dnmt1b* and is involved with transcription factor binding and chromatin modification.

Figure 2.10 shows the key for the components of secondary protein structure, alpha helixes which usually span membranes and is concerned with the binding of DNA, beta strands, coils and turns.

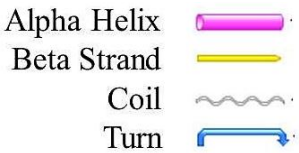


Figure 2.10: Secondary structure key. An identification for the structures present in the secondary structure of proteins.

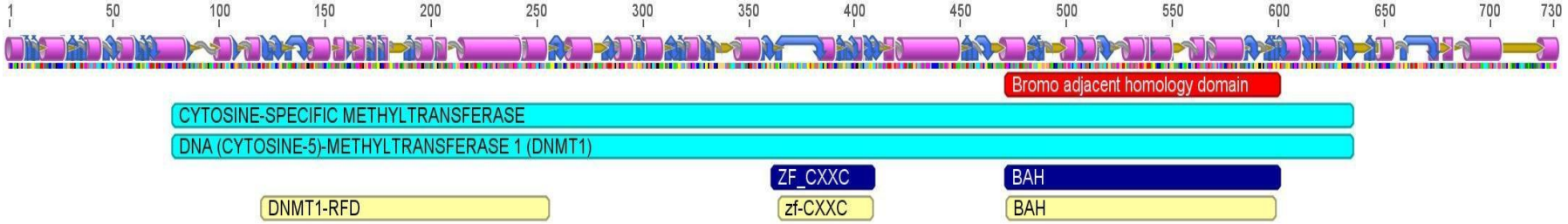


Figure 2.11: Secondary structure & conserved domains in *Dnmt 1a*. Dnmt1a contains a cytosine-specific methyltransferase domain, a DNA (cytosine-5)-methyltransferase 1 (Dnmt1) domain, a cytosine specific DNA methyltransferase replication foci domain, multiple BAH domains and CXXC zinc finger domains.

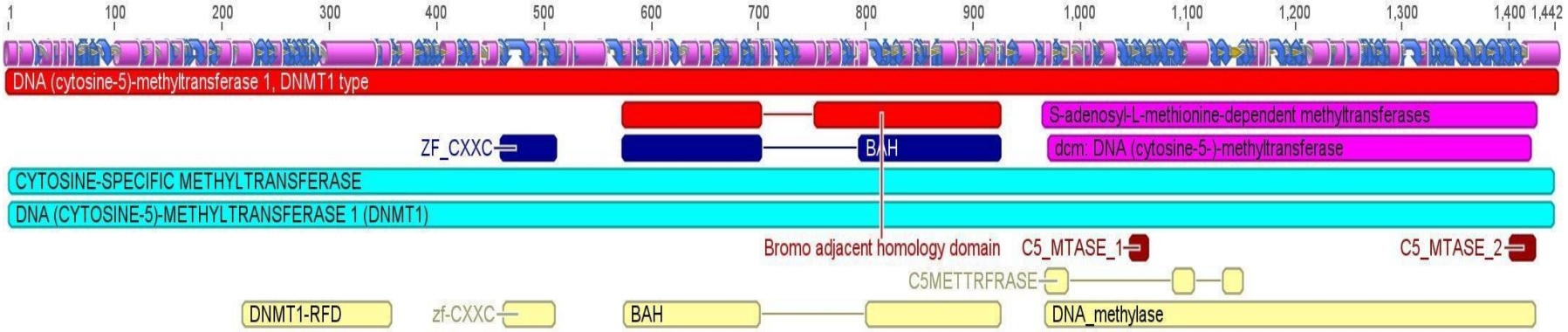


Figure 2.12: Secondary structure & conserved domains in *Dnmt1b*. In addition to the domains found in *Dnmt1a, Dnmt1b* also has an additional DNA-cytosine methyltransferase domain, a C-5 cytosine-specific DNA methylases C- terminal signature, a cytosine-specific DNA methyltransferase signature (PR00105), a S-adenosyl-L-methionine-dependent methyltransferases, a C-5 cytosine-specific DNA methylase and a DNA (cytosine-5)-methyltransferase 1, DNMT1 type domain.

Dnmt2 shares a majority of its conserved domains with other methyltransferases (Figure 3.13). It includes a C-5 cytosine-specific DNA methylase (PF00145) domain and a S-adenosyl-L-methionine-dependent methyltransferase domain (SSF53335) like *Dnmt1b* and *Dnmt3*. It also shares 3 domains in common with Dnmt1b- a cytosine-specific DNA methyltransferase signature (PR00105), a C-5 cytosine-specific DNA methylases C-terminal signature (PS00095) and the DNA-cytosine methyltransferase domain (TIGR00675).

BtDnmt3 also shares a few domains with other methyltransferases- the S-adenosyl-L-methionine-dependent methyltransferase domain (SSF53335), C5_MTASE_1 **(**PS00094) and the C-5 cytosine-specific DNA methylase (PF00145) domain. In addition, it also has a domain from an unnamed family (PTHR23068). The PWWP domain (PS50812), is specific to *Dnmt3* and is composed of 100 to 150 amino acids (Figure 3.14). However, its precise function is unknown (Chen *et al.*, 2004). Tudor/PWWP/MBT (SSF63748) is a superfamily connected with DNA replication and repair. Finally, *Dnmt3* contains a FYVE/PHD zinc finger superfamily (SSF57903)-members include a phosphatidylinositol-3-phosphate binding domain, a PHD domain and a variant PHD-like domain. It too is involved with DNA replication and repair.

MBD proteins consist of an unidentified family (PTHR12396) and a methyl-CpG binding domain (PF01429) (Figure 2.15). It attaches to DNA which possesses symmetrically methylated CpGs (Nan, 1993).

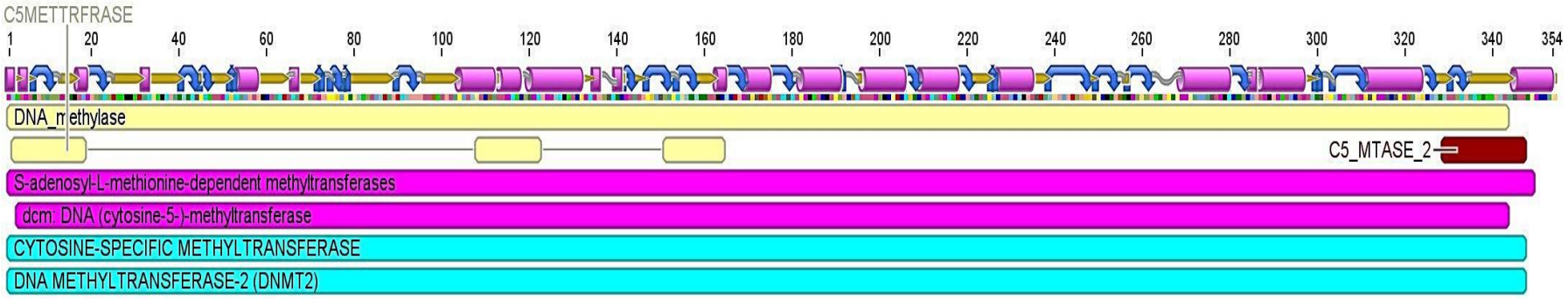


Figure 2.13: Secondary structure & conserved domains in *Dnmt2* includes a C-5 cytosine-specific DNA methylase domain, a S-adenosyl-L-methionine-dependent methyltransferase domain, a cytosine-specific DNA methyltransferase signature, a C-5 cytosine-specific DNA methylases C-terminal signature and the DNA-cytosine methyltransferase domain.

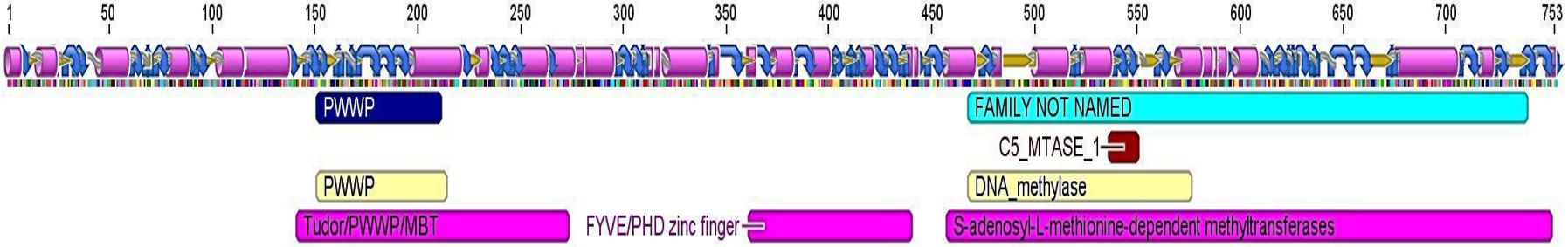


Figure 2.14: Secondary structure & conserved domains in *Dnmt3*. *BtDnmt3* has a S-adenosyl-L-methionine-dependent methyltransferase domain), **C5_MTASE_1,** a C-5 cytosine-specific DNA methylase domain, a domain from an unnamed family, a PWWP domain and a FYVE/PHD zinc finger superfamily of domains.

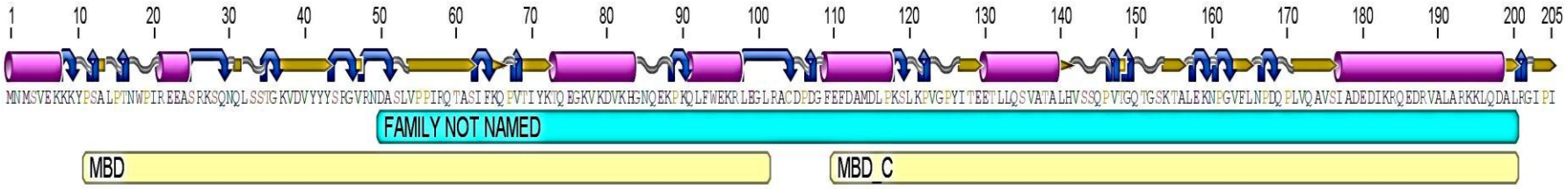


Figure 2.15: Secondary structure & conserved domains in *MBDP.* *B. terrestris* MBD proteins consist of an unidentified family of domains and a methyl-CpG binding domain.

2.4 Discussion

2.4.1 Summary of results

The aim of this chapter was to establish the presence of a methylation system in *Bombus terrestris* and a prediction was made that *B. terrestris*, like *Apis mellifera*, has a full complement of methyltransferases. This prediction was correct, with *B. terrestris* exhibiting a full complement of methyltransferases as well as a methyl-binding protein. The process of amplifying *B. terrestris* methylation genes using primers made with *A. mellifera* methylation gene sequences resulted in the complete sequencing of *BtDnmt2* and the partial sequencing of *BtDnmt1b*, *BtDnmt3* and *BtMBP*. These partial sequences were then used to make probes to scan the *B. terrestris* BAC library, resulting in the positive hybridisation of 68 BAC library clones. The publication of partially annotated *B. terrestris* and *B. impatiens* genomes helped complete the partial sequences found by PCR and BAC library screening. The published *Bombus* genomes also gave an opportunity to assess the structure and the evolution of bumblebee methyltransferases in relation to other insect methyltransferases. The sequence analysis reflected the bees' phylogeny, with *B. terrestris* and *B. impatiens* sequences showing high conservation. The bumblebee methylation genes were more conserved with A. *mellifera* compared to those of *Nasonia vitripennis.* An analysis of the conserved domains in each *B. terrestris* methyltransferase confirmed the presence of domains usually found in vertebrate and invertebrate methyltransferases.

Given the high level of conservation within the *B. terrestris* and *A. mellifera* genomes, it is surprising that *Dnmt2* was the only methyltransferase gene which was conserved to a level which enabled successful primer construction and complete amplification. One

possible solution to aid in the construction of successful primers would be to use only the *A. mellifera* methyltransferase sequences when making primers, instead of adding other insect sequences to the alignment. For example, when making a primer set to amplify *B. terrestris* Dnmt3, not using *A. pisum* in an alignment with *N. vitripennis* and *A. mellifera* would have decreased the amount of variation that was found, enabling the primers to be more specific.

2.4.2 Implications and further work

Cytosine methylation involves the transfer of a methyl group from S-adenosyl methionine to cytosine, a process catalysed by methyltransferases. The *B. terrestris* genome, like human and *A. mellifera* genomes, contains a full complement of methyltransferases. It is yet unknown if these methyltransferases will act similarly to those found in other organisms; *BtDnmt1* may be a maintenance methyltransferase, *BtDnmt2* a RNA cytosine methyltransferase, *BtDnmt3* may be involved in establishing novel methylation markers and *BtMBD* may play a role in maintenance and *de novo* methylation. *Dnmt3* is overexpressed in reproducing *A. mellifera* workers, suggesting a connection with methylation in germ lines (Cardoen *et al.*, 2011). However, functional studies of the *B. terrestris* methylation machinery, such as an analysis of the tertiary structure and the presence of binding sites, as well as performing enzyme assays, would aid in the characterisation of the methyltransferase genes.

Information about the role of methyltransferases in insect-specific biological functions is very limited and makes exploring the exact role of methylation in genomic imprinting

and phenotypic plasticity in social insects problematic. However, it is possible to comparatively assess what is currently known about insect methyltransferases. *BtDnmt1a* and *Dnmt1b* exhibit the classical structures found in other *Dnmt1* genes throughout the animal kingdom- the methyltransferase and bromo-adjacent homology (BAH) domains. *BtDnmt2* surprisingly seems to have a cytosine specific methyltransferase domain, even though it acts on RNA instead of DNA. *BtDnmt3*, like *BtDnmt1a* and *Dnmt1b* exhibit a conserved domain found in other *Dnmt3* genes- the Pro-Trp-Trp-Pro (PWWP) domain (Schaefer & Lyko, 2007). BAH domains are involved in protein-protein interactions and result in regulating transcription (Callebaut *et al.*, 1999). The function of the PWWP domain is not clear. Studies have shown that in human *Dnmt3b2*, the domain's structure allows the methyltransferase to bind with DNA. In addition, this domain is not crucial for *in vitro* CpG methylation (Qui *et al.*, 2002). Until recently it was assumed that, in addition to insects not having a standard compliment of methyltransferases, there wasn't a link between the evolutionary relationship of two species and the number of methyltransferases they possessed (Lyko & Maleszeka, 2011). For example, *A. mellifera* and *N. vitripennis* are both hymenopteran species, yet possess a different number of *Dnmt1* orthologues (Lyko & Maleszeka, 2011). However, evolutionary distance between species may not be the reason for insects showing a diverse range of methyltransferase compliment, as it has been shown *B. terrestris* and *A. mellifera* have an identical compliment of methyltransferases. Insect methylation systems appear to reflect the diversity of insect biological processes.

The *B. terrestris* methylation system shares a lot of similarities with the methylation system of the pea aphid, *Acyrthosiphon pisum*. Both *ApDnmt1a* and *ApDnmt1b* and

BtDnmt1a and *BtDnmt1b* contain a BAH conserved domain, which is characteristic of *Dnmt1* genes. However, *ApDnmt1a* and *ApDnmt1b* are 77% identical, as compared to the 27% of *BtDnmt1a* and *BtDnmt1b*. The similarity of the methylation systems which is the presence of a full complement of methyltransferases, observed in Hymenopteran species *B. terrestris*, *A. mellifera* and *N. vitripennis* suggest that methylation can have a regulatory role in Hymenopteran genomes, possibly by regulating phenotypic plasticity during the development of different castes (Lyko & Maleszeka, 2011).

As mentioned previously, although the *Bombus* and *Apis* genera have independently evolved for 100 million years, a great percentage of their genomes and even full chromosomes are highly conserved (Stolle *et al.*, 2011). More intra and inter chromosomal shuffling has occurred in other taxa in the same time as the divergence between *B. terrestris* and *A. mellifera*. For example, extensive chromosome shuffling in homologous chromosome arms and deviated genomes in *D. melanogaster* and *D. erecta* has been observed; these two species diverged approximately 10 million years ago (mya). Extensive chromosomal organisation has also been observed in mammalian species that evolved independently in a relatively short time of 40-90 mya. Thus, this high level of homology in *B. terrestris* and *A. mellifera* is unexpected. This kind of homology is usually present in closely related species, such as rats and mice, which diverged only 16 mya or even that found within a species of *Apidae- A. mellifera* and *A. florea* which split 20-25 mya (Stolle *et al.*, 2011). This high level of genomic conservation exists despite both species having a high genome recombination rate, higher than that for other insect or vertebrate species (Stolle *et al.*, 2011).

Two splice variants of *AmBMDP* exist. *AmMBDP-1* was mostly active *in vivo* and is similar to vertebrate *BMDP3. In vitro* function binding capacity of this family varies but *in vivo* are important parts of chromatin remodelling complexes (Wang *et al.*, 2006). It would be interesting to see if the case is the same in *B. terrestris*, as chromatin remodelling is a process resulting in genomic imprinting.

Further work that can be conducted on the methylation system in *B. terrestris* includes calculating of the CpG ratio and quantification of 5mC in the *B terrestris* genome, possibly using LC-ESI-MS/MS (liquid chromatography/electrospray ionisation mass spectrometry). DNA is first enzymatically hydrolysed by using three enzymes in tandem. The products are then sequentially divided using a variation of high performance liquid chromatography (Song *et al.*, 2005). In addition, a better understanding of the expression of *Bombus* methyltransferases would be invaluable for caste and life stage transcriptomic studies. The expression of the methyltransferases in a range of tissues such as the antennae, brain, thorax and ovaries can be assessed in various life stages and castes.

Chapter 3 : Methylated genes in *Bombus terrestris*

3.1 Introduction

The evolution of sterile castes in social insect systems is considered to be a key evolutionary shift (Maynard Smith & Szathmáry, 1995). With the evolution of sterile castes, individuals formed a new level of biological cooperation and resulted in a higher level of organisation, resulting in a reproductive division of labour. Females in social insect societies develop to be morphologically and behaviourally specialised either as a reproducing queen or a non-reproducing worker (Pereboom *et al.*, 2005). This differentiation depends more on the individual's environment during the process of caste differentiation than genotypic differentiation. Thus, caste differentiation of the reproducing queens and non-reproducing workers would be expected to be due to differentially expressed genes which are shared by both female castes.

Understanding how differential gene expression leads to caste determination in social Hymenoptera is crucial as the information would demonstrate the exact nature of the association between gene expression and phenotypic diversity. A previous study showed that genes expressed in queens and non-reproducing workers at the larval and adult stages were distinct (Pereboom *et al.*, 2005). In addition, while *Bombus terrestris* and *Apis mellifera* share some genes associated with caste differentiation, the expression patterns varied. For example, queen destined larvae in *B. terrestris* showed upregulated gene expression during the early stages of development, while the same genes were upregulated in worker-destined larvae later on in development. Hence, increased expression of same genes at different times of development demonstrates that the timing of gene expression is vital in *B. terrestris* caste differentiation. Also, experiments looking into differential expression of genes in reproducing and non-

reproducing *B. terrestris* workers have found that non-reproducing workers have higher metabolic rates compared to reproducing workers of the same age (Cardoen *et al.*, 2011). The differential expression of genes in workers with different reproductive abilities is in accord with Bourke (1988), which states that reproducing workers would not risk their reproductive output by engaging in risky behaviours such as foraging.

Methylation of genes can lead to differential expression. Genes in *A. mellifera* can either have a high CpG content or a low CpG content as a result of DNA methylation. This varied level of CpG content can be explained by the role that methylation seems to play in *A. mellifera.* Genes rich in CpG dinucleotides, which are hypomethylated in germlines, are associated with development, while those with a low CpG dinucleotide content, which are hypermethylated in germlines, are associated with cellular processes such as translation, transcription and the cell cycle (Wang & Leung, 2009). However, not all caste-specific genes were rich in CpG dinucleotides- some of the metabolism genes, while low in CpGs, were also seen to be differentially expressed across castes (Elango *et al.*, 2009). Another study found that nearly a third of the annotated genes involved with 'housekeeping' functions in *A. mellifera* are methylated at CpG dinucleotides in the intra-genic regions (Foret *et al.*, 2009).

Reproductive plasticity has been observed in *B. terrestris* workers (Alaux *et al.*, 2007). Essentially, while workers can have mature ovaries early on in the colony's life (Duchateau & Velthuis, 1989), they abstain from laying eggs while the queen is present. However, if the colony loses its queen and becomes queen-less, the workers start to lay eggs after a week (Alaux *et al.*, 2007). When queen-less workers of the same species

were kept together, a dominance hierarchy was attained (Alaux *et al.*, 2006a) and the oldest worker laid eggs. The younger workers' oocytes were resorbed. In order to test this, mini-colonies of a single 2 day old worker and two 1 day old workers (from a different donor colony) were established. When the bees were dissected after 7 days, the older worker was always the egg layer. On occasions when the mini-colony was allowed to continue for two weeks, an increase in ovary development in the dominant workers was not observed (Alaux *et al.*, 2007). As mentioned in Chapter 1, methylation helps an organism, population or a species adapt to a changing or new environment, thereby avoiding extinction. In addition to exploring methylated genes in castes and life stages in *B. terrestris* as an element of analysing genomic imprinting, methylated genes can also be studied as a means of calculating the level of plasticity *B. terrestris* is capable of.

Prediction:

The methyltransferase genes in *Bombus terrestris* are functional and differentially methylated genes will be found in different castes (queens, non-reproducing workers, queen-right reproducing workers and queen-less reproducing workers) and life stages (larval stages, pupae and callow workers).

3.2 Methods

The experiments in this chapter assess whether the *Bombus terrestris* methylation machinery found in Chapter 2 are functional, which would result in methylated genes. In addition, the methylation profiles of different caste and life stages will also be analysed, thus observing how the bee's environment (absence or presence of a queen which in turn determines reproductive state) influences genetic plasticity.

Two methods were used to assess the methylation status of genes in *B. terrestris.* First, a set of loci was analysed using bisulphite sequencing. Secondly, a behavioural experiment was carried out to promote queen-less egg laying in workers and the methylation profiles of a range of life stages and caste was analysed, using restriction enzymes.

3.2.1 Bee keeping

The bees used for the techniques in Section 3.2 were reared according to Section 2.2.1.

3.2.2 Bisulphite conversion of DNA to detect methylation

Bisulphite conversion of DNA results in the deamination of cytosine to uracil. The converted DNA is then amplified with PCR and sequenced. The methylation status of the cytosines is analysed by comparing the original DNA sequence with the treated DNA sequence. Methylation is detected when a cytosine in the original sequence is unchanged in the treated DNA sequence, while non-methylated cytosines appear as

thymine, because these cytosines were converted to uracil during the treatment. Pre-treating DNA with sodium bisulphite is a quick, cost effective application which allows a nucleotide-level assessment of methylation at a locus (Komori *et al.*, 2011).

Bisulphite sequencing was conducted on casein kinase β 2 and a protein coding locus, both of which were found to be methylated in *A. mellifera* (Wang *et al.*, 2006). DNA was extracted using the Promega Wizard® SV Genomic Purification System. The EZ DNA Methylation-Gold Kit™ (Zymo Research) was used in the modification of genomic DNA. The manufacturer's instructions were followed.

Primers for both bisulphite treated and untreated DNA were designed using Primer3, version 0.4.0 (http://frodo.wi.mit.edu/) on default settings. They were designed to amplify two overlapping sections of DNA, covering the promoter, as well as the gene sequence. Additional primers for treated DNA were designed using MethPrimer (http://www.urogene.org/methprimer/) (Table 3.1). Stock solutions were made to 100 pmol/µl, with working solutions being 10 pmol/µl.

PCRs were optimized according to section 2.3.1, by varying annealing time and temperature. The products were then visualised according to section 2.2.5 and PCR purification and sequencing was conducted according to section 2.2.6.

Table 3.1: Bisulphite sequencing primers. Primers used for assessing methylation status of the casein kinase 2 β gene. (Primers in red were used on bisulphite treated DNA).

NAME	T^m	SEQUENCE	SIZE
CK2B.1-F	64.0	TACATCATTGGGACGAGCAA	1121
CK2B.1-R	63.5	AACCATTGCTTCTCCAGGAA	
CK2B.2-F	63.7	CCGGACTAAATGAACAAGTACCA	886
CK2B.2-R	63.5	TTTCAGCGCCTTGCTAAAAT	
CK2B.M1-F	57.7	GTGGTTAATTAAAGTATTAAAAGGTAAGT	987
CK2B.M1-R	61.7	AAATAACCATCCAATTCATCATCTAC	
CK2B.M2-F	52.8	TAAATGAATAAGTATTATATTATAGGTAAG	868
CK2B.M2-R	53.1	ACTAAAATATATAAAAATCAATATAAAAA	

3.2.3 Restriction analysis of DNA to detect methylation

3.2.3.1 Behavioural assay

Since differential gene methylation of *B. terrestris* queens, queen-less reproducing workers and control non-reproducing workers (from a queen-right colony) was being assessed, a queen-less environment which promoted worker reproduction was needed. Reproductive plasticity was induced according to *Alaux et al.* (2007).

Newly emerged *B. terrestris* workers were collected from three colonies twice a day. They were tagged and 2 coloured spots were placed either side of their abdomen in a specific manner in order to differentiate the day each bee was tagged. The age of the worker bee was distinguished according to the colour of the numbered tag and the colours of the two spots of nail varnish. Each bee was placed in a plastic pot and cooled on ice for 20 minutes. Using a wooden toothpick, a drop of gum was placed on the thorax. A numbered Opalith tag (Christian Graze KG, Germany) was then positioned (Figure 3.1). The bees were placed back in their natal colony 20 minutes later once the gum had dried.

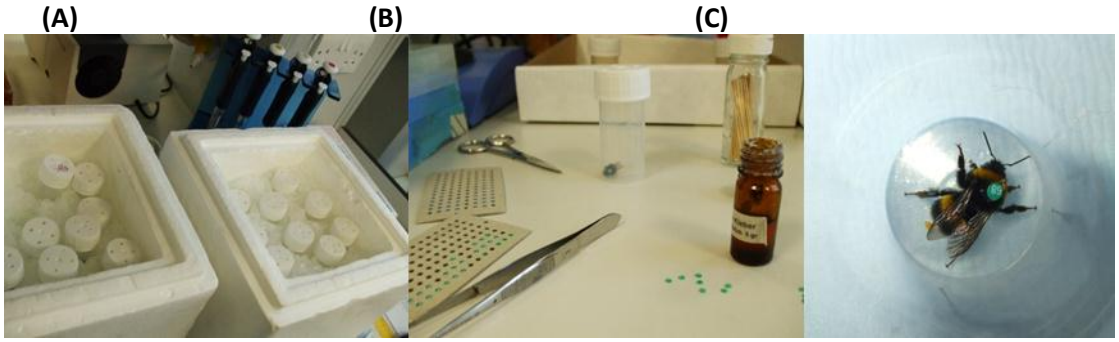


Figure 3.1: Tagging bees. (A) The bees were kept on ice (B) The coloured & numbered tags and gum used (C) A tagged bee.

Mini-colonies were made two days later, where each marked bee was placed in a smaller clear Perspex box (17cm x 5.5cm x 11.5cm) with two one-day old workers from different donor colonies. A total of 35 mini-colonies were made. A blend of pollen and honey was provided as substrate for cell construction. According to Alaux *et al.* (2007), the older worker becomes the egg layer. The older bee from each mini-colony was collected seven days later, snap frozen in liquid nitrogen and stored at -80°C. Non-reproducing bees of the same age (10 days) were collected from natal colonies for the control sample of non-reproducing workers.

Each of the three natal colonies was observed daily to establish when the competition point was reached. Competition point was achieved if workers were observed laying eggs or destroying queen laid eggs (Cnaani *et al*., 1997). Only one of the colonies reached completion point and one egg laying worker was collected (sample R1). Larvae were collected from natal colonies 1 and 2. Larval stages were determined using weight, according to Table 3.2 (Cnaani *et al.*, 1997), snap frozen in liquid nitrogen and stored at -80°C.

Table 3.2: Categorising larval stages. The maximum weight for each larval instar in *B. terrestris* (Cnaani *et al.*, 1997).

INSTAR	MAXIMUM WEIGHT (g)
1	0.006
2	0.039
3	0.095
4	0.375

3.2.3.2 Extraction of genomic DNA using phenol:chloroform

NTE was prepared according to the table 3.3 (Wang *et al*., 2006).

Table 3.3: NTE preparation. Reagents and volume added to make 250ml of NTE

COMPONENT	FINAL CONC.	FOR 250ml OF NTE
NaCl	100mM	25ml
Tris	50mM	12.5ml
EDTA	10mM	2.5ml
SDS	1%	12.5ml
Triton X-100	0.01%	25µl
H2O	-	197.5ml

Each adult sample was composed of three legs and half of the abdomen per, while each larval sample was composed of half a larva. 700µl of NTE and 17.5µl of Proteinase K (20mg/ml; Sigma-Aldrich) was added to each sample and incubated on a heat block at 50°C for 3 hours. One volume of phenol:chloroform (Fisher Scientific) was added and spun in a centrifuge (Progen) for 10 minutes at 10,000g. The upper phase was collected and treated with 1µl RNAse (10mg/ml) for 10 minutes. Another phenol:chloroform extraction step was performed. The samples were precipitated with two volumes of absolute ethanol, incubated at room temperature for 20 minutes and then centrifuged at maximum speed (13,000 RPM) for 15 minutes. All of the supernatant was discarded

and the tubes were dried in the fume hood. The sides of the tube were then washed with 50µl of distilled water to elute the DNA.

3.2.2.3 Amplification of Intermethylated Sites (AIMS)

AIMS is a process where genomic DNA is digested by a pair of restriction endonucleases, one which is sensitive to methylation and the other which is insensitive to methylation, to produce a methylation profile composed of loci with a methylated cytosine at each end (Frigola *et al.*, 2002).

One of two pairs of restriction enzymes can be used for each AIMS procedure. *SmaI*/*XmaI* recognises CCCGGG sites. The second isoschizomer pair *HpaII*/*MspI* recognises CCGG sites and is generally used for genomes with a low CpG content, such as *A. mellifera*.

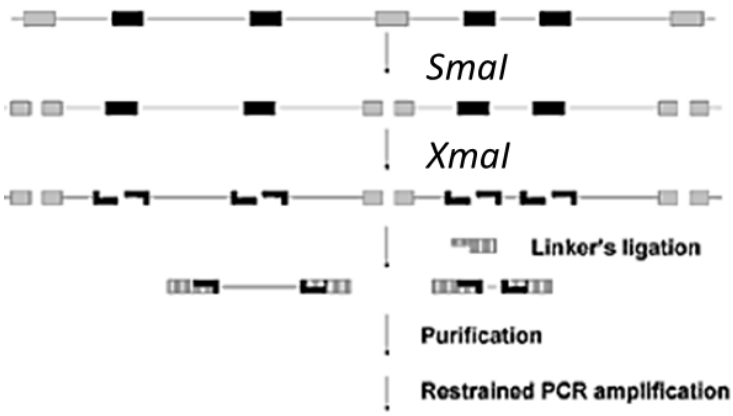


Figure 3.2: AIMS. The first line shows genomic DNA (solid line) and 7 CCCGGG sites, three which are non-methylated (grey boxes) and four which are methylated (black boxes). *SmaI* cuts the non-methylated sites, leaving blunt ends. *XmaI* then cuts the methylated sites, leaving sticky ends (CCGGG overhang). Adaptors are ligated to the sticky ends (line four) and all fragments which have an adaptor at each end are amplified using adaptor specific primers. (Frigola *et al.*, 2002)

Thus, when using *SmaI*/*XmaI*, the methylation sensitive *SmaI* leaves a blunt end in unmethylated DNA (cleave results in CCC/GGG). The methylated sites remain unchanged. Methylated DNA is cleaved at specific sites with *XmaI*, which results in sticky ends (C/CCGGG). Adaptors are added to only the sticky ends and the methylated fragments are amplified using adaptor specific primers (Figure 3.2). These primers are composed of the adaptors, followed by 1-3 nucleotides chosen arbitrarily. Once amplified, the methylated DNA fragments can be visualised on a gel. Each fragment is composed of a locus which is flanked by a methylated cytosine (Jordá *et al.*, 2009). In the case of *HpaII*/*MspI*, the recognition site CCGG is cleaved at C/CGG, with *HpaII* sensitive to methylation and *MspI* being insensitive to methylation.

The preparation of adaptors (50μM) was conducted according to Frigola *et al.* (2002), where 25μl of the oligonucleotide Blue (100μM) (5'–ATTCGCAAAGCTCTGA -3') and 25μl of the oligonucleotide MCF (100μM) (5'-CCGGTCAGAGCTTTGCGAAT-3') were incubated at 65°C for 2 minutes and then cooled to room temperature over 1 hour in a thermocycler. For the second set of restriction enzymes, adaptors were constructed as above using ADPT B (100μM) (5'-ATTCGCAAAGCTCTGA-3') and ADPT B-CG (100μM) (5'-P-CGTCAGAGCTTTGCGAAT-3').

The DNA was first digested with the methylation sensitive restriction endonuclease *SmaI*. 1μg of DNA was digested with 1.5μl of NEB4 10x Buffer, 0.1μl *SmaI* and 3.4μl ddH_2O to make up a 15μl reaction that was incubated for 1 hour at 25°C. All reactions were incubated in a thermocycler. The entire product was then digested for another hour

at 37°C with1.0µl of NEB4 10x Buffer, 0.5µl of *Xma*I, 0.5µl of bovine serum albumin (BSA) and 8µl of ddH_2O for a final volume of 25µl (Jorda *et al.*, 2009).

A separate set of digestions was conducted using *HpaII*/*MspI*, which targets a different set of nucleotides compared to *SmaI*/*XmaI*. The DNA was digested with the methylation sensitive restriction endonuclease *HpaII*. 1µg of DNA was digested with 1.5µl of NEB4 10x Buffer, 1.0µl *HpaII* and 2.5 ddH_2O to make up a 15µl reaction that was incubated for 1 hour at 37°C. *HpaII* ends were filled in using 1.0µl of Polymerase I, Large Klenow fragment (NEB) and 1.5µl of dNTP (500µM). This reaction was incubated for 15 minutes at 37°C, and inactivated by incubating for 10 minutes at 75°C. The above product was then digested for another hour at 37°C with 0.75µl of NEB4 10x Buffer, 0.5µl of *MspI* and 6.25µl of ddH_2O for a final volume of 25µl (Jorda *et al.*, 2009).

Ligation of the sticky ends of digested DNA was conducted with 25µl of digested DNA, 8µl of T4 buffer, 2µl of T4 ligase and 20µl of adaptor (2nmol), which was incubated for 10 minutes at room temperature (20°C-25°C). This was followed by inactivation of the enzyme by incubating the samples for 10 minutes at 65°C (Jorda *et al.*, 2009). Each digestion/ligation reaction was aliquoted and stored in the freezer. A few aliquots were initially stored at -20˚C but they degraded. Hence, all subsequent aliquots were stored at -80˚C.

The amplification of sequences with the adaptors was conducted using the primer sets A, B, C and D for DNA digested with *SmaI/XmaI*. iiA1, iiA2, iiB and iiC were used to amplify DNA digested with *HpaII/MspI* (Table 3.4).

Table 3.4: AIMS primer information. Primer sequences used to amplify digested DNA

Primer Set	Primer	Sequence
Set A	A1	BLUE-CCGGG-CTA
	A2	BLUE-CCGGG-TGG
Set B	B1	BLUE-CCGGG-CTG
	A2	BLUE-CCGGG-TGG
Set C	C1	BLUE-CCGGG-CGCG
	C2	BLUE-CCGGG-CAAC

Reaction volumes of 25µl were composed of 12.5µl of 2x reaction buffer (Sigma), 1µl of each primer, 3µl of DNA, 0.5µl of 10mM $MgCl_2$ and 7µl of water. For Primer set A and B, the PCR cycle consisted of 30 2-step cycles (15s at 94°C and 1.5 minutes at 74°C). PCR with Primer set C consisted of 30 3-step cycles (15s at 94°C, 45s at 68°C and 1minute at 74°C). The PCR cycles were preceded with a denaturing step of 95°C for I minute and ended with an extension of 72°C for 5 minutes. Amplification of *MspI/HpaII* products for all 3 sets of primers consisted of 35 3-step cycles (1 minute at 94°C, 45s at 63°C and 1.5 minutes at 72°C). The PCR cycles were preceded with a denaturing step of 95°C for 2 minute and ended with an extension of 72°C for 5 minutes.

The PCR products were separated by electrophoresis on a 3% agarose gel stained with EtBr at 55V for 1.5 hours, with 1µl of loading buffer with 5µl of sample. The DNA

bands were visualised under UV light on a transilluminator. Once the presence of products were verified, 10µl of PCR product was mixed with 3µl of 1x loading buffer (Elchrom) and run on 9% poly(NAT) gels (Elchrom) using the Origins electrophoresis system (Elchrom) at 120V for 81 minutes at 55°C. 1x TAE was used as a buffer and was changed every 3 runs as with accordance to the manufacturer's notes. The gel tank was connected to a power supply using a Biometra powerpack. The gel was then peeled off the plastic backing and stained in the dark with SybrGold (1:10000 dilution in TAE) in a staining tray on a rocking platform for 40 minutes. The gel was destained with 100ml TAE, again in the dark on the rocking platform for 30 minutes. The gel was visualised under UV light on a transilluminator.

The bands on each gel lane were compared, with bands not uniformly present in castes or life stages isolated using the BandPick™ (Elchrom). The bands were then re-amplified using the original primer pair and PCR mixture, electrophoresed on a 1% agarose gel and excised. The MinElute PCR Purification Spin Protocol was used in the purification of the successful PCR products. The NanoDrop 1000 spectrophotometer was used to quantify the PCR products, with 1.5µl of sample was loaded onto the pedestal. The results were analysed using the software provided (NanaDrop 1000 v3.7).

3.2.2.4 Cloning of AIMS products

The cloning steps were carried out in a sterile environment using a flow hood (MicroFlow, UK). Ampicillin was stored in stock concentrations of 100mg/ml at -20°C. X-gal, for blue and white selection, was stored as a stock solution of 20mg/ml at -20°C and used at 40µg/ml. IPTG stock solution of 0.1M was stored at 4°C. The amount of

PCR product to be used was calculated using the equation below, with a 3:1 insert: vector molar ratio.

$$\text{ng of insert} = \frac{50\text{ng vector} \times \text{kb insert}}{3.0\text{kb vector}} \text{ x } \frac{3}{1}$$

LB medium was made using 10g Bacto®-Tryptone, 5g Bacto®-Yeast Extract and 5g NaCl per litre. The pH was adjusted to 7.0 before autoclaving. LB plates were made with ampicillin, where, 15g of Bioagar (BioGene) was added to each litre of LB medium. Again, the pH was adjusted to 7 and autoclaved. The medium was cooled to approximately 50°C and ampicillin was added to a final concentration of 100µg/ml. The medium was then poured into 85mm petri dishes in sterile conditions. The agar was left to harden for 30 minutes and plates were then stored at 4°C. 100µl of 100mM IPTG and 20µl of 50mg/ml X-Gal were then spread over the surface of each LB-ampicillin plate, followed by incubation at 37°C in order for the IPTG and X-gal to be absorbed. PCR products were cloned using the pGEM®-T Easy Vector System (Promega).

3.2.2.5 Colony PCR, glycerol stocks and isolating plasmid DNA from E. coli cultures

Blue-white selection is a technique which detects the successful ligation of a DNA insert into a vector. When DNA is ligated into a plasmid, the α peptide is disrupted and β-galactosidase is not produced. Thus, cells transformed with plasmids containing an insert produce white colonies, while cells with a plasmid which do not have an insert produce blue colonies, allowing for the identification of cells which were successfully ligated (Ullmann *et al.*, 1967).

Colony PCRs were conducted to screen clones which were either completely white or had a blue centre. Inserts were checked using the M13- F & R sequencing primers (CGCCAGGGTTTTCCCAGTCACGAC and TCACACAGGAAACAGCTATGAC) as well as the primers used with the original PCR the fragment came from. A negative control of a blank plasmid was used. An additional 10 minute denaturing step was added to the beginning of the PCR programme to lyse the bacteria.

If the colony PCR indicated that the insert was of the expected size, the colony was removed from the reference plate with a sterile toothpick and was incubated overnight in a shaking incubator at 37˚C in a 15ml falcon tube of 5ml LB and 1.5µl ampicillin. Glycerol stocks were made in 1.5ml screw top Eppendorf tubes by adding 700µl of the bacteria solution to an equal volume of 30% glycerol solution. Each tube was gently inverted to mix the contents, flash frozen in liquid nitrogen and stored at -80˚C. Plasmids were isolated using the GenElute^{TM} Plasmid Miniprep Kit (Sigma-Aldrich) and sent to Source Bioscience (Nottingham) for commercial sequencing (Level 2).

3.2.2.6 Sequence analysis

Sequences were analysed using Finch TV and Geneious (v5.5). Vector contamination in the cloned sequences was analysed using Geneious. The sequences were also analysed using BLASTx against the *B. terrestris* genome. If the result was a hypothetical protein, another BLASTx was carried out using the *A. mellifera* and *D. melanogaster* genomes. The GO website was used to classify the proteins., while KEGG (Kyoto Encyclopaedia of Genes and Genomes) was used to annotate the pathway each sequence was a part of.

3.3 Results

3.3.1 Bisulpite sequencing

Table 3.5 illustrates the annealing and extension temperatures used to amplify the casein kinase β 2 gene.

Table 3.5: Primers beginning with 'CK' were used to amplify the casein kinase β 2 gene while primers those with 'PC' were used to amplify the protein coding locus. Primers in red were used on bisulphite treated DNA. T^A and T^E with '-'denote primers that did not work.

PRIMER	SEQUENCE	T^A/˚C	T^E/˚C
CK2B.1-F	TACATCATTGGGACGACAA	55	72
CK2B.1-R	AACCATTGCTTCTCCAGGAA		
CK2B.2-F	CCGGACTAAATGAACAAGTACCA	55	70
CK2B.2-R	TTTCAGCGCCTTGCTAAAAT		
CK2B.M1-F	GTGGTTAATTAAAGTATTAAAAGGTAAGT	54	72
CK2B.M1-R	AAATAACCATCCAATTCATCATCTAC		
CK2B.M2-F	TAAATGAATAAGTATTATATTATAGGTAAG	49	72
CK2B.M2-R	ACTAAAATATATAAAAATCAATATAAAAA		
CK2B.3-F	TACGAATCCTGCGTCTTTT	57	74
CK2B.3-R	ATGGCCATCCAATTCATCAT		
PC1.1-F	TGAGAGATGTATCGCTTTATG	-	-
PC1.1-R	TTCTTGTTCAGAGGCCAACCTC		
PC1.2-F	GAGGTTGGCCTGAACAAGAA	-	-
PC1.2-R	ACCTGGACCAAGTGGTTCAC		
PC1.3-F	CTTGTGAACCACTTCGTCCAG	57	74
PC1.3-R	AGCATTTCCAACAATATTGCCTA		

Figure 3.3 shows the results for primer pairs CK2B.3 and PC3. Primer pair PC3 produced multiple products, while CK2B.3 was sequenced.

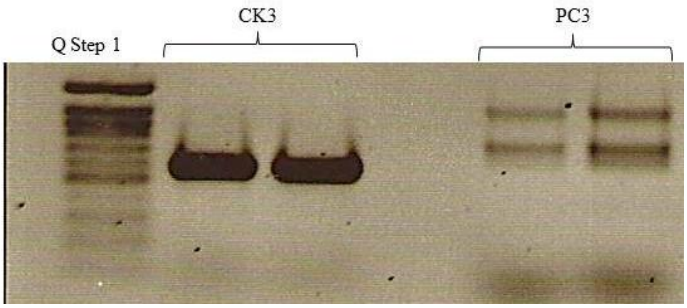


Figure 3.3: Gel of 2 separate reactions using primer sets CK3 and PC3

The bisulphite PCRs were not always reproducible and occasionally produced multiple products. For example, figure 3.4 shows the results for 4 PCRs, with both CK2B.CM1 and CK2B.CM2. A master mix was made for each primer set and then split into two separate reactions, which were run on the same amplification programme and then run on separate lanes on a gel. As the gel shows, only one of the CK2B.CM2 reactions worked, while one of the CK2B.CM1 reactions has amplified less product than the other. Multiple products were regularly produced.

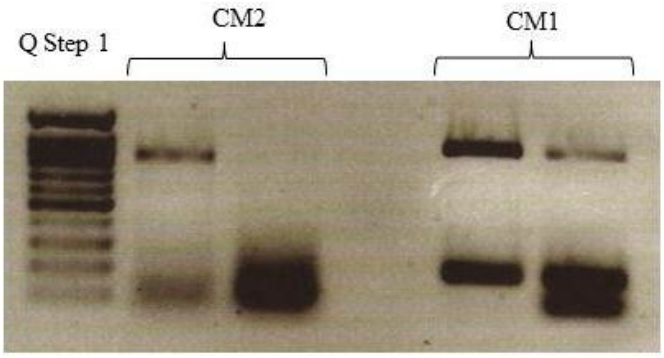


Figure 3.4: Gel of 2 separate reaction using primer sets CK2B.CM1 and CK2B.CM2.

In addition, when the correct product was sent through a DNA clean up column, there was insufficient DNA for sequencing. On the few occasions that bisulphite treated samples were sent to be sequenced, the results were of poor quality. Therefore, the initial sequence and the bisulphite-treated sequence could not be aligned to assess cytosine methylation.

3.3.2 Amplification of inter-methylated sites (AIMS)

Although 35 mini-colonies were established, eggs were laid by the dominant worker in only 15 mini-colonies. The samples were coded from A-G, Q and R to simplify marking gel lanes. ‘A’ were 10 day old reproducing workers, ‘B’ were 10 day old non-reproducing workers, ‘Q’ were queens, ‘C’ were callow workers, ‘D’ were pupae, ‘E’ were L4 larvae, ‘F’ were L3 larvae and ‘G’ were L2 larvae. A total number of 5

replicates were conducted. Primer Set A was run on all 5 replicates, while primer sets B and C were run on replicates 1-4. A total of 13 AIMS gels were run. Table 3.6 gives a summary of the samples used for AIMS analysis.

While a total of 136 bands were picked, only 51 were analysed. A number of bands did not successfully amplify, produced different product sizes when re-amplified, could not be ligated successfully into a plasmid. Finally, a number of sequences were not of a good quality and could not be run on BLAST.

Table 3.6: Samples used in AIMS. The restriction enzyme digest was conducted on five sets of samples. Only two replicates (replicate 2 and 3) had a queen, as one of the three queens in natal colonies died. A=10 day old reproducing workers, B= 10 day old non-reproducing workers, C= callow workers, D= pupae, E=L4 larvae, F= L3 larvae, G= L2 larvae and Q= queens.

REPLICATE	NUMBER OF SAMPLES								
	Q	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**R**
Replicate 1	0	3	3	3	3	2	2	2	0
Replicate 2	1	3	3	3	3	2	2	2	0
Replicate 3	1	3	3	3	3	2	2	2	0
Replicate 4	0	3	3	3	3	1	1	2	0
Replicate 5	0	3	3	3	3	0	1	2	1

Figure 3.5 shows the mass and distribution of larval instars for colony 1 and 2 respectively. Colony 1 produced 6 L2 larvae (G), 7 L3 (F) larvae, and 3 L4 (E) larvae, while 6 L2 larvae, 3 L3 larvae and 4 L4 stage larvae were collected from colony 2.

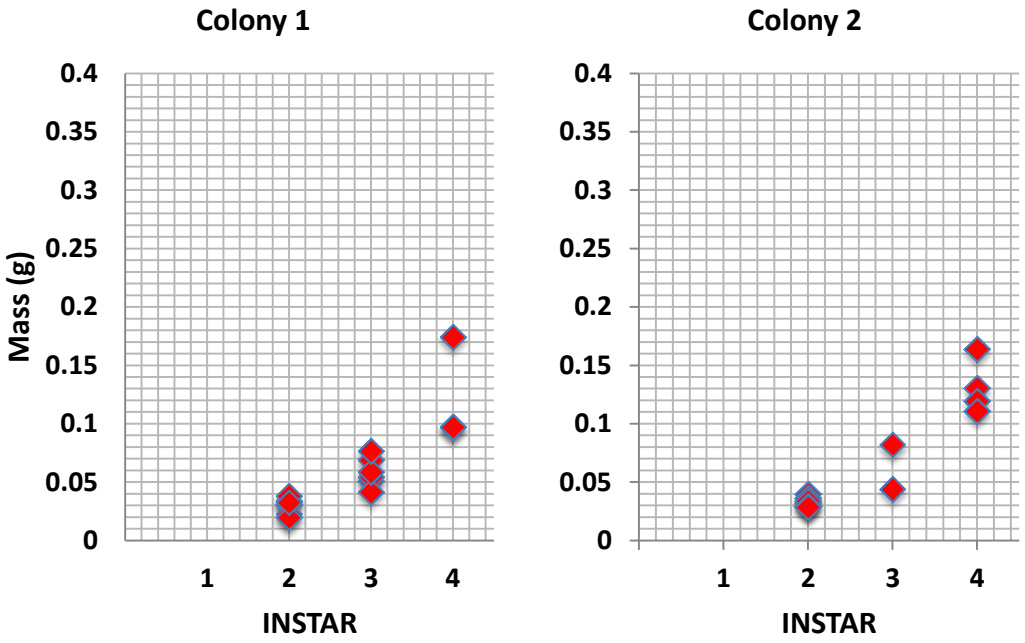


Figure 3.5: Larval weights and instar for Colony 1 and Colony 2 samples.

Only sections of 3 gels are included in the results section (See appendix for the remaining gels). Figure 3.6 shows two gels, replicate 1 and 5 for primer set B. Bands appearing on one sample and not the other for the same primer set were concluded to be hypermethylation (Frigola *et al.*, 2002). Decreases in the intensity were thought to be hypomethylation. Bands which exhibited hypermethylation or hypomethylation were picked out and re-amplified are shown in red.

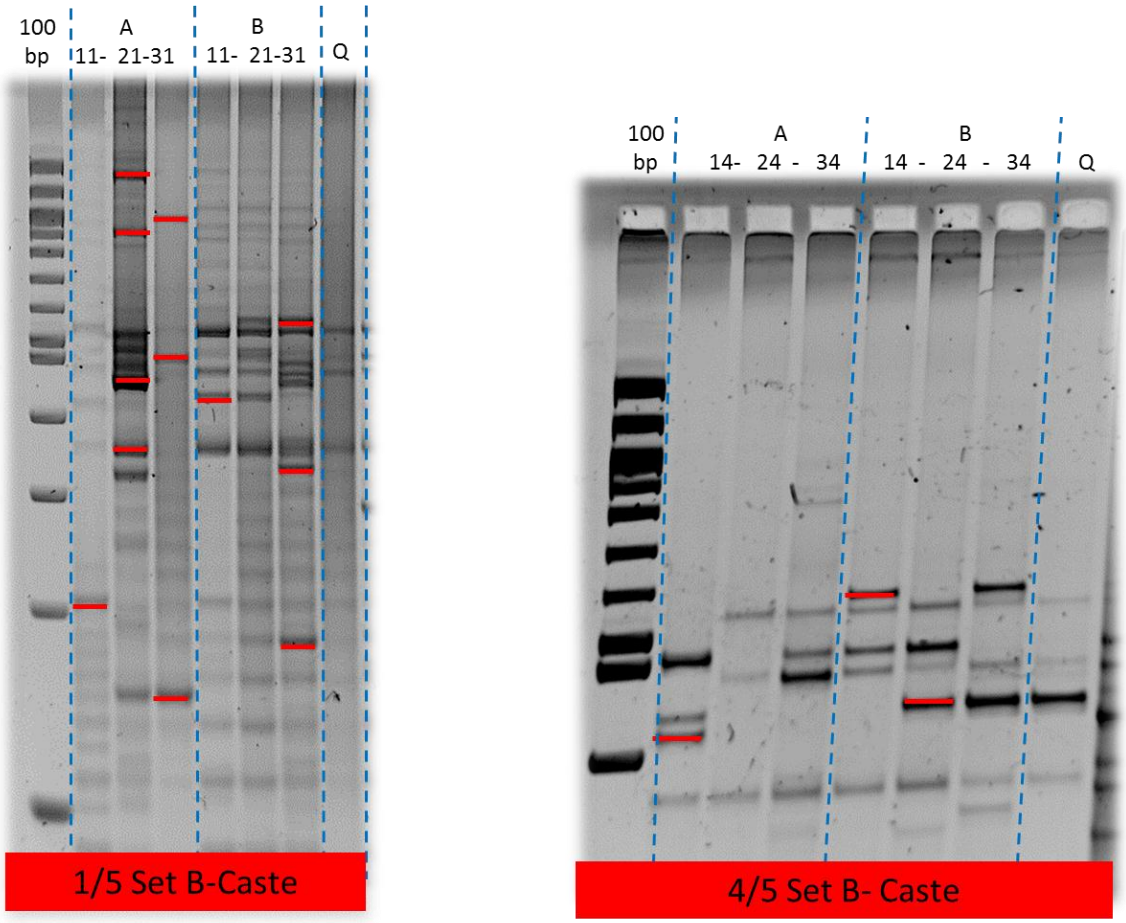


Figure 3.6: AIMS gels for replicate 1 and 4, using primer set B. A=10 day old reproducing workers, B= 10 day old non-reproducing workers and Q= queens. Samples were labelled according to sample type, colony number (second number) and sample number (first number). Thus, sample 21 would be the first sample from colony 2, while sample 31 would be the 1st sample from colony 3. The bands which were picked, amplified and cloned are highlighted in red.

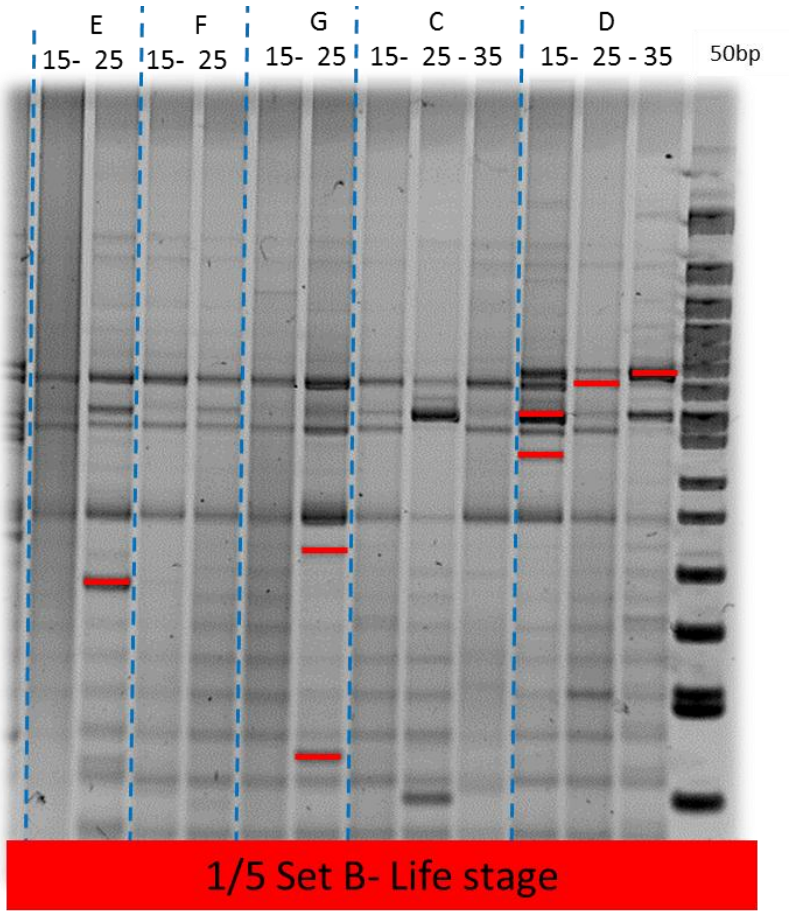


Figure 3.7: AIMS gel for replicate 1 using primer set B. Samples used were life stages. C= callow workers, D= pupae, E=L4 larvae, F= L3 larvae and G= L2 larvae. Samples were labelled according to sample type, colony number (second number) and sample number (first number). Thus, sample 25 would be the fifth sample from colony 2, while sample 15 would be the 5th sample from colony 1. The bands which were picked, amplified and cloned are highlighted in red.

Figure 3.7 shows a gel for replicate 1, primer set B. This gel looked at the methylation profiles for life stages. Again, bands of interest which were picked out are shown in red.

Table 3.7: AIMS results. BLAST analysis of bands which were sequenced.

BAND	DETAILS	*B. terrestris (e value)*	*A. mellifera*	*D. melanogaster*
1	Hypermethylated in 10 day old non-reproducing workers	long-chain fatty acid transport protein 4-like XP_003402267.1 (2e-13)		
2	Hypermethylated in 10 day old non-reproducing workers	hypothetical protein LOC100644011 XP_003395259.1 (0.004)	LOW QUALITY PROTEIN: hypothetical protein LOC412081 XP_395547.4 (8e-04)	bifid, isoform A NP_525070.2 (4e-04)
3	Hypermethylated in L4 larvae	acyl-CoA synthetase short-chain family member 3, mitochondrial-like XP_003398692.1 (2e-06)		
4	Hypermethylated in L3 larvae	hypothetical protein LOC100648629 XP_003395625.1 (0.001)	hypothetical protein LOC409729 XP_393227.4 (6e-04)	CG15021 NP_647902.1 (5e-04)
5	Found in all	collagen alpha-1(IV) chain-like XP_003400654.1 (7e-04)		
6	Found in all	hypothetical protein LOC100648629 XP_003395625.1 (0.001)		
7	Found in all	hypothetical protein LOC100648629 XP_003395625.1 (0.001)	TPA: TPA_inf: troponin H isoform 2 DAA05524.1 (6e-04)	
8	Hypermethylated in callow workers	long-chain fatty acid transport protein 4-like XP_003394923.1 (5e-07)		
9	Hypermethylated in pupae	long-chain fatty acid transport protein 4-like XP_003402267.1 (3e-07)		

BAND	DETAILS	*B. terrestris (e value)*	*A. mellifera*	*D. melanogaster*
10	Hypermethylated in pupae	hypothetical protein LOC100643403 XP_003395581.1 (8e-04)	RING finger protein nhl-1-like XP_391967.4 (0.001)	
11	Found in all	hypothetical protein LOC100646265 isoform 2 XP_003398440.1 (4e-04)	histamine-gated chloride channel 1 precursor NP_001071279.1 (6e-04)	
36	Hypomethylted in 10 day old non-reproducing workers	hypothetical protein XP_003394397.1 (2e-05)	DNA replication licensing factor MCM8 XP_395500.4 (1e-07)	
38	Hypomethylated in 10 day old non-reproducing workers	PREDICTED: protein takeout-like XP_003398008.1 (0.002)	general transcription factor 3C polypeptide 1-like XP_003249863.1 (2e-04)	
39	Hypermethylated in 10 day old reproducing workers	glucose dehydrogenase [acceptor]-like XP_003402779.1 (1e-22)		
42	Hypermethylated in 10 day old non-reproducing workers	hypothetical protein LOC100651274 XP_003393638.1 (0.003)	hypothetical protein LOC412077 XP_395543.4 (0.003)	RE01762p ACX83592.1 (8e-04)
44	Hypermethylated in life stages	zinc finger protein 287-like XP_003395208.1 (2e-04)		
45	Only found in L2 larvae	hypothetical protein LOC100644923 XP_003394859.1 (0.001)	e3 ubiquitin-protein ligase MIB2-like XP_392821.4 (0.001)	
49	Hypermethylated in pupae	glucose dehydrogenase [acceptor]-like XP_003403320.1 (6e-14)		
50	Found in all	long-chain fatty acid transport protein 4-like XP_003402267.1 (9e-20)		

BAND	DETAILS	*B. terrestris (e value)*	*A. mellifera*	*D. melanogaster*
51	Hypermethylated in pupae	long-chain fatty acid transport protein 4-like XP_003402267.1 (2e-20)		
53	Hypermethylated in 10 day old reproducing workers	hypothetical protein LOC100651712 XP_003393880.1 (3e-05)		
54	Hypermethylated in 10 day old reproducing workers	mitogen-activated protein kinase kinase kinase kinase 4-like isoform 3 XP_003403323.1 (0.001)		
57	Hypermethylated in 10 day old reproducing workers	enolase-like XP_003402548.1 (0.002)	multiple inositol polyphosphate phosphatase 1-like, partial XP_001121719.2 (2e-04)	
58	Found in all	sterol regulatory element-binding protein cleavage-activating protein-like XP_003403180.1 (1e-04)		
59	Hypermethylated in 10 day old non-reproducing workers	ubiquitin carboxyl-terminal hydrolase 34-like XP_003395465.1 (3e-04)		
61	Hypermethylated in 10 day old non-reproducing workers	enolase-like XP_003402548.1 (0.001)		
62	Found in all	peroxidasin-like XP_003396953.1 (0.001)		
63	Found in all	enolase-like XP_003402548.1 (0.001)		

BAND	DETAILS	*B. terrestris (e value)*	*A. mellifera*	*D. melanogaster*
71	Hypermethylated in 10 day old reproducing workers	homocysteine-responsive endoplasmic reticulum-resident ubiquitin-like domain member 2 protein-like XP_003400729.1 (5e-04)		
73	Hypermethylated in life stages	LOW QUALITY PROTEIN: segmentation protein paired-like XP_003400641.1 (0.005)		
74	Hypermethylated in life stages	hypothetical protein LOC100642227 XP_003401227.1 (1e-04)	hypothetical protein LOC552100 XP_624482.3 (0.001)	LD05615p AAQ22553.1 (0.002)
75	Hypermethylated in life stages	26S proteasome non-ATPase regulatory subunit 4-like XP_003401112.1 (0.004)		
79	Hypermethylated in L2 larvae	microtubule-associated serine/threonine-protein kinase 4-like isoform 1 XP_003393083.1 (2e-04)		
81	Hypomethylated in 10 day old non-reproducing workers	facilitated trehalose transporter Tret1-like XP_003399625.1 (1e-04)		
83	Hypermethylated in 10 day old non-reproducing workers	long-chain fatty acid transport protein 4-like XP_003402267.1 (1e-20)		
87	Only found in queens	long-chain fatty acid transport protein 4-like XP_003402267.1 (7e-19)		
94	Hypermethylated in life stages	hypothetical protein LOC100650284 XP_003397181.1 (0.001)	hypothetical protein LOC408830 isoform 1 XP_392361.4 (0.001)	RH42429p AAM29596.1 (1e-04)
97	Hypermethylated in L2 larvae	poly(rC)-binding protein 3-like XP_003401437.1		

BAND	DETAILS	*B. terrestris (e value)*	*A. mellifera*	*D. melanogaster*
98	Hypermethylated in L2 larvae	venom dipeptidyl peptidase 4-like XP_003399213.1 (0.001)		CG9392 | NP_649136.1 (8e-05)
103	Hypermethylated in pupae	hypothetical protein LOC100643257 XP_003400177.1 (6e-04)	hypothetical protein LOC552005 XP_624389.2 (6e-04)	
110	Hypermethylated in 10 day old reproducing workers	hypothetical protein LOC100646088 XP_003403143.1 (3e-04)	2-aminoethanethiol dioxygenase-like XP_001120557.1 (6e-04)	
111	Hypermethylated in 10 day old non-reproducing workers	hypothetical protein LOC100650777 XP_003398477.1 (2e-04)	hypothetical protein LOC409497 XP_393004.3 (2e-04)	klarsicht, isoform D NP_001097463.1 (4e-04)
115	Found in all	long-chain fatty acid transport protein 4-like XP_003402267.1 (3e-20)		
116	Found in all	acyl-CoA synthetase short-chain family member 3, mitochondrial-like XP_003398692.1 (3e-11)		
119	Hypermethylated in 10 day old reproducing workers	long-chain fatty acid transport protein 4-like XP_003402267.1 (1e-20)		
120	Hypermethylated in 10 day old reproducing workers	long-chain fatty acid transport protein 4-like XP_003402267.1 (7e-15)		
121	Hypermethylated in 10 day old reproducing workers	acyl-CoA synthetase short-chain family member 3, mitochondrial-like XP_003398692.1 (1e-10)		

BAND	DETAILS	*B. terrestris (e value)*	*A. mellifera*	*D. melanogaster*
122	Hypermethylated in 10 day old reproducing workers	transcriptional activator cubitus interruptus-like XP_003395140.1 (8e-04)	Golgi apparatus protein 1-like isoform 1 XP_397171.3 (0.002)	
123	Hypermethylated in 10 day old reproducing workers	long-chain fatty acid transport protein 4-like XP_003402267.1 (2e-06)		
130	Hypermethylated in queen-right reproducing worker	UPF0047 protein yjbQ-like XP_003403368.1 (4e-05)	UPF0047 protein yjbQ-like XP_394840.2 (5e-04)	
132	Hypermethylated in queen-right reproducing worker	hypothetical protein LOC100649672 XP_003401370.1 (0.004)	hypothetical protein LOC552558 XP_624937.1 (0.005)	TPA: TPA_inf: HDC00957 DAA03524.1 (2e-05)

Table 3.7 tabulates the BLAST results from the AIMS bands which were successfully cloned and sequenced. Certain sequences, like enolase occurred multiple times. Table 3.8 shows further analysis of the results in Table 3.7, including the biological processes and pathways they are involved in.

Table 3.8: AIMS results. Analysis of AIMS results using Gene Ontology & KEGG. If the biological process was not found in KEGG it is within "( )".

GENE	GO BIOLOGICAL PROCESS	KEGG PATHWAYS
long-chain fatty acid transport protein 4-like	Transport	-Fatty acid metabolism, -Adipocytokine signalling -Cutin, suberine and wax biosynthesis
acyl-CoA synthetase short-chain family member 3, mitochondrial-like	Enzyme	-Fatty acid metabolism -Peroxisome biogenesis
collagen alpha-1(IV) chain-like	Cellular component	-ECM-receptor interaction
PREDICTED: protein takeout-like	Binding	-Unknown
glucose dehydrogenase [acceptor]-like	Enzyme	- Pentose phosphate pathway
zinc finger protein 287-like	Nucleic acid binding (DNA)	-Unknown
mitogen-activated protein kinase kinase kinase kinase 4-like isoform 3	Enzyme	- MAPK signalling - GnRH signalling - Toll-like receptor signalling - Fc epsilon RI signalling
enolase-like	Binding	-Glycolysis / Gluconeogenesis -Methane metabolism -RNA degradation
sterol regulatory element-binding protein cleavage-activating protein-like	Structural protein	-Unknown
ubiquitin carboxyl-terminal hydrolase 34-like	Binding	(Regulation of gene silencing)
peroxidasin-like	Binding	(Plays a role in extracellular matrix consolidation, phagocytosis and defense)

GENE	GO BIOLOGICAL PROCESS	KEGG PATHWAYS
homocysteine-responsive endoplasmic reticulum-resident ubiquitin-like domain member 2 protein-like	Membrane protein	-Protein processing in endoplasmic reticulum
LOW QUALITY PROTEIN: segmentation protein paired-like	Nucleic acid binding (DNA)	-Protein processing in endoplasmic reticulum
26S proteasome non-ATPase regulatory subunit 4-like	Protein Binding	-Proteasome
microtubule-associated serine/threonine-protein kinase 4-like isoform 1	ATP binding	-Unknown
facilitated trehalose transporter Tret1-like	Transport	-Unknown
poly(rC)-binding protein 3-like	Nucleic acid binding (RNA)	-Unknown
venom dipeptidyl peptidase 4-like XP_003399213.1 (0.001)	Unknown	(Protein digestion and absorption)
transcriptional activator cubitus interruptus-like	Nucleic acid binding (DNA)	Hedgehog signaling pathway
UPF0047 protein yjbQ-like	Unknown	-Unknown

Figure 3.8 separates the results according to their functions. A majority of proteins had unknown functions or were hypothetical proteins. Next, many proteins were parts of binding complexes. A few were either structural proteins, or parts of cellular membranes, while others were enzymes.

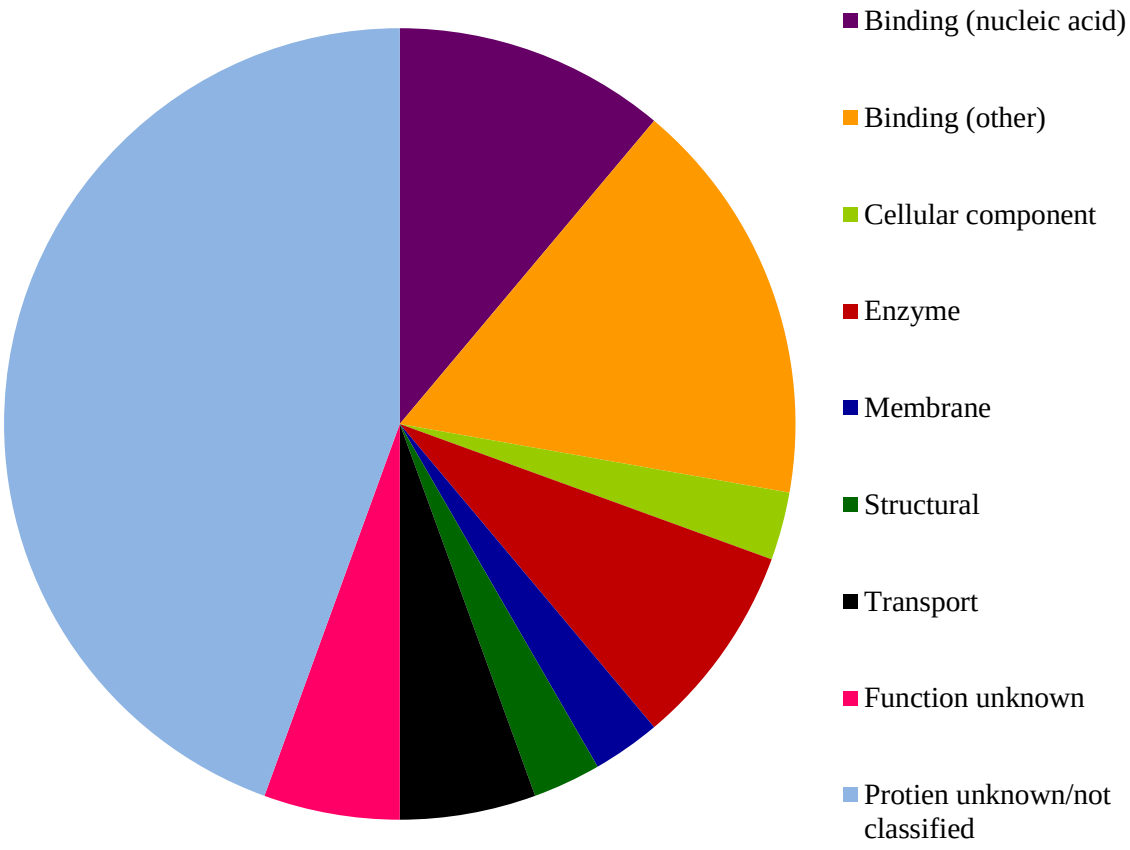


Figure 3.8: Chart representing the proteins in table 3.7 and table 3.8. The pie chart showing the proportion of the proteins according to their GO functions

3.4 Discussion

3.4.1 Summary of results

The aim of this chapter was to assess if the *Bombus terrestris* methylation system is functional and a prediction was made that differentially methylated genes will be found in different castes (queens, non-reproducing workers, queen-right reproducing workers and queen-less reproducing workers) and life stages (larval stages, pupae and callow

workers). This prediction was correct, with *B. terrestris* exhibiting differentially methylated genes in castes and life stages.

The research in this chapter was conducted on two different sets of data. The samples used for bisulphite sequencing were workers from a queen-right colony (colony where the queen was present). The samples used in the AIMS analysis were from a queen-right colony and queen-less mini-colonies. The non-reproducing controls were from a donor colony with a queen, while the reproducing workers were from the mini-colonies. It was ensured that the corresponding control for each reproducing eldest worker from the mini-colony was from the same natal colony. *Apis mellifera* has been documented to have a low CpG content and it was assumed that *B. terrestris* would be similar, thus making the *MspII*/*HpaI* isoschizmer pair more successful. This was not the case however and *MspII*/*HpaI* digested DNA resulted in methylation profiles of poor quality. Thus, the experiment was conducted using *SmaI*/*XmaI*. The methylation profiles generated from AIMS led to a considerable number of loci being sequenced. 44.4% of the genes found were proteins which have not been confirmed to be expressed *in vivo,* or hypothetical proteins (Sivashankari1 & Shanmughave, 2006). While the corresponding sequences were found in *A. mellifera* or *Drosophila melanogaster*, a better understanding of the function of these hypothetical proteins would be possible when the *B. terrestris* genome is fully annotated.

The poor performance of bisulphite treatment was disappointing, especially since the problem seemed to be with the sequencing. One possible alternative would be to conduct whole genome bisulphite sequencing (WGBS) for methylation analysis or

using methylation specific microarrays. While such techniques are expensive and require extensive bioinformatical analysis, they would be a comprehensive indicator of global methylation.

3.4.2 Implications and further work

A physiological difference between queens and non-reproducing workers is the accumulation of glycogen and fat in their fat body. Workers with underdeveloped ovaries also possess smaller fat bodies (Duchateau & Velthuis, 1989). Long-chain fatty acid transport protein and acyl-CoA synthetase short-chain proteins are both components in the metabolism of fatty acids and were found to be differentially methylated in reproducing and non-reproducing workers. It is possible that these two genes are differentially expressed in reproducing and non-reproducing workers due to egg laying. In queens however, accumulation of glycogen aids with overwintering. As stated by Pereboom *et al.* (2005), caste differentiation is possibly due to differentially expressed genes. A few genes found by AIMS- long-chain fatty acid transport protein 4-like, acyl-CoA synthetase short-chain family member 3, mitochondrial-like and enolase-like genes- were found to be methylated at different places in caste and life stages. Thus, these genes may be differentially expressed according to the caste and age of the individual. Enolase has been found to be overexpressed in non-reproducing *A. mellifera* workers in queen-less colonies compared to reproducing workers (Cardeon *et al.*, 2011).

On another note, AIMS usually performs poorly in insects (Walsh *et al.*, 2010). The poor performance has been attributed to the fact that the enzymes target mammalian DNA methylation patterns. Thus, a more comprehensive methylation profile may be constructed if a combination of restriction enzymes were used to target more commonly found insect sequences. In order to further evaluate the function of methylation in *B. terrestris,* the AIMS experiment could be modified to use localised tissue for each study group. For example, methylation profiles in a range of tissues such as the antennae, brain, thorax and ovaries in various life stages such as reproducing and non-reproducing workers and queens could be assessed because DNA methylation patterns are capable of being tissue specific (Wang *et al.*, 2006). Embryos would be another good tissue to run AIMS on because sexual differentiation occurs within a few hours of eggs being laid (Bigot *et al.*, 2010). A similar study conducted by Bigot *et al.* (2010) also found that methylation was involved in development and caste. Embryos were highly methylated while methylation decreased at alternative developmental stages. In addition, Bigot *et al.* (2010) found that a transposable element, *BotmarI* experienced many rearrangements and deletions during sexual differentiation and development.

One of the genes found to be differentially methylated in this study was acyl-CoA synthetase. It is a component of the juvenile hormone (JH) synthesis pathway, a hormone crucial for the development of many insects. For example, JH is involved in the regulation of parthenogenic, sexual, apterous and winged morphs in *A. pisum* (Walsh *et al.*, 2010). Like in *B. terrestris,* a JH pathway component was also found to be methylated in *Acyrthosiphon pisum*. JH is associated with the development and maintenance of castes in *A. mellifera* (Barchuk *et al.*, 2002). However, JH does not influence oogenesis or the activation of ovaries in adult *A. mellifera* workers (Cardoen

et al., 2011). Alternatively, JH and vitellogenin influences the switch from nurses to foragers in workers (Guidugli *et al.,* 2005).

Kucharski *et al.* (2008) also compared global gene expression in *A. mellifera Dnmt3* knock out and control individuals. This showed a range of differentially expressed genes which were involved in the transport of lipids, regulation of hormones, post-translational modification, ribosomal biosynthesis and novel genes. Approximately 6000 conserved and universally expressed genes in *A. mellifera* have been found to be methylated (Lyko *et al.*, 2010). In addition, levels of methylation are high near alternatively spliced exons, suggesting that this methylation results in the regulation of transcript variants (Lyko & Maleszeka, 2011). Evidence for this proposal is apparent in the way that differential methylation appears to regulate the splicing of *GB18602*, a putative membrane protein. *GB18602* can result in two transcripts, one which codes for a complete protein and a shorter isoform created by the addition of a cassette-exon which contains a premature stop codon (Lyko *et al.*, 2010). The DNA which codes for the cassette-exon is richly methylated in workers but not in queens. This indicates that the differential methylation results in a caste-specific transcriptional event. The short variant is expressed more in the queen brain than in the worker brain, illustrating that the methylation dependent splicing of DNA controls caste-specific protein suppression (Lyko & Maleszeka, 2011).

There are numerous alternative reasons for exploring methylation in *B. terrestris* other than to confirm the presence of a platform for genomic imprinting. *B. terrestris* is a member of a social insect species important to the agricultural industry, as bumblebees

forage for longer periods and ranges, in worse weather and visited more flowers per minute compared to *A. mellifera*. Bumblebees also pollinate more fruit and vegetable flowers than honeybees (Willmer *et al*., 1994). The agricultural industry has been affected by decreasing bumblebee populations, which have been blamed on disease, loss of habitat and pesticides. Methylation of esterase genes in the peach potato aphid *Myzus persicae* results in the aphids developing a resistance to insecticides (Field *et al*., 2004). A similar mechanism of methylation-induced insecticide resistance may be found in *B. terrestris*, aiding agriculture at a time when global population is causing a strain on food availability. Decreases in *B. terrestris* populations due to loss of habitat may be counteracted by phenotypic plasticity, as phenotypic plasticity can accelerate evolutionary adaptation (Espinosa-Soto *et al*., 2011). The study of decreasing bumblebee populations can be aided with a knowledge of a biological system which could help bumblebees cope better with changes to their environment or with the colonisation of a new environment.

This thesis explores phenotypic plasticity and genomic imprinting as a result of methylation and therefore AIMS, which provides profiles for CpG methylation, was used to find evidence of a functioning methylation system. However, *D. melanogaster* and the pea aphid *A. pisum*, have shown CpA methylation. If methylation in *B. terrestris* is to be characterised, addressing the possible existence of non-cytosine based methylation needs to be considered.

Chapter 4 : Worker reproduction & genomic imprinting in *Bombus terrestris*

4.1 Introduction

Imprinted genes are usually genes which regulate embryonic growth and development (Queller, 2003). The conflict theory predicts that paternally imprinted genes increase growth, while maternally imprinted genes decrease growth (Burt & Trivers, 2006). Hymenopterans are haplodiploid- females develop from fertilised eggs and males from unfertilised eggs, leading to relatedness asymmetries. Other facets of relatedness are caused by queens in some species like *Apis mellifera* being polyandrous or by colonies having multiple queens. In addition, resource allocation interactions is not only between the mother and her offspring, but also involves sisters, cousins and half-brothers (Queller, 2003). These aspects of hymenopteran societies make them a good starting point to independently test the role of genomic imprinting in Haig's conflict theory.

Imprinted genes have not yet been found in any social insect system. However, it is accepted that the methylation machinery necessary is present in many species (Kronfrost *et al.*, 2008). As discussed earlier, there is evidence to suggest that DNA methylation is used to store epigenetic information in honeybees. Kucharski *et al.* (2008) illustrated that knocking out *Dnmt3* leads to an increase in the number of larvae which develop into queens. Another example of potential imprinting in an insect is the defensive behaviour portrayed by Africanised honeybees which has been suggested to be correlated with paternally imprinted genes (Guzman-Novoa *et al.*, 2005). A study conducted on the rove beetle *Aleochara bilineata* shows that the solitary parasitoid larvae not only avoid superparasitising full siblings, but also cousins who they are related to via their father. The same behaviour is not shown to cousins related by the mother and this maternal cousin level of avoidance is the same as that for non-related

individuals (Lizé *et al.*, 2007). The examples of *A. mellifera* defensive behaviour and ovary development in *Dnmt3* knock outs as well as rove beetle larval behaviour demonstrate that imprinting is feasible in some insects.

Genomic imprinting occurs in tissues involved with the transfer of resources from parent to offspring, such as the placenta in mammals or the endosperm in angiosperms. Workers in social insect societies provide food for the young larvae, even though they themselves rarely reproduce. If the workers start to reproduce, it will be at a cost to their siblings, the queen-laid larvae as the amount of care provided for the queen-laid larvae decreases. Thus, genes for worker reproduction would be involved with the manipulation of resources for young and therefore have a potential to be imprinted.

Many predictions for imprinting in Hymenopteran genes have been made (Queller 2003). At each locus, the fitness effects of imprinting for males, workers and queens, as well as the probability of the gamete being present in the individual has to be accounted for (Kronauer, 2008). From a worker's perspective, their unfertilised eggs can only result in haploid males. Therefore, the expected expression pattern should increase the male's inclusive fitness (Kronauer, 2008). It has been shown that in species where workers are more related to the queen's sons than other worker's sons, fewer males are produced (Wenseleers & Ratnieks, 2006).

A prediction was made in Chapter 1 that genes in workers which promote reproduction in queen-right *Bombus terrestris* colonies, due to matrigenes and patrigenes being

subjected to different selectional pressures, matrigenes would be imprinted while the patrigene would be expressed. However, this study explores the imprinting of genes associated with worker reproduction in queen-less colonies. (Figure 4.1, area within the dotted lines). Thus, a patrigene has a 50% chance of being present in a worker's son or her sister's son, while a matrigene has a 50% chance of being in her own son but only a 25% chance of being present in another worker's son. It would be expected that the patrigene will be imprinted (not expressed) and the matrigene will be expressed in genes that promote worker reproduction in queen-less *B. terrestris* colonies. The converse is true for genes which hamper worker reproduction, where it would be expected that the patrigene will be expressed while the matrigene will be imprinted.

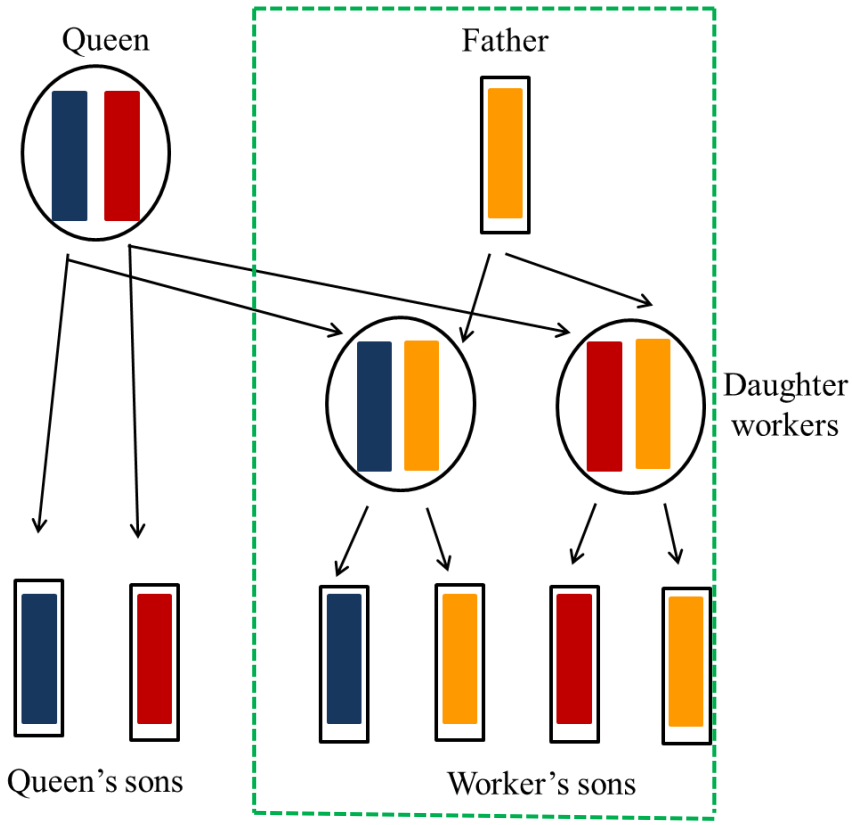


Figure 4.1: Genomic imprinting in *B. terrestris* queen-less colonies. Spread of maternally derived and paternally derived alleles in *B. terrestris* in queen-right colonies and queen-less colonies (within the green dotted lines)

Prediction:

The paternal allele will be imprinted (not expressed) and the maternal allele will be expressed in genes that promote reproduction in *B. terrestris* workers from queen-less colonies.

4.2 Methods

The aim of this chapter was to Test Haig's theory using worker reproduction in queen-less *Bombus terrestris* colonies.

First, alleles were genotyped by locating single nucleotide polymorphisms (SNPs) in exons of candidate genes for worker reproduction found from a literature search and Chapter 3. The experimental method used in this chapter to identify SNPs is single strand confirmation polymorphism (SSCP), where DNA is denatured and electrophoresed on a GMA gel. The electrophoretic mobility of the single stranded DNA is determined by the nucleotide sequence- the sequences form hairpins, pseudoknots and triple helices according to the bases (Nielsen *et al*, 1995) and the size of the strand. If a SNP is present, the mobility of the DNA will be different in each allele (Figure 4.2) and will be reflected in the gel.

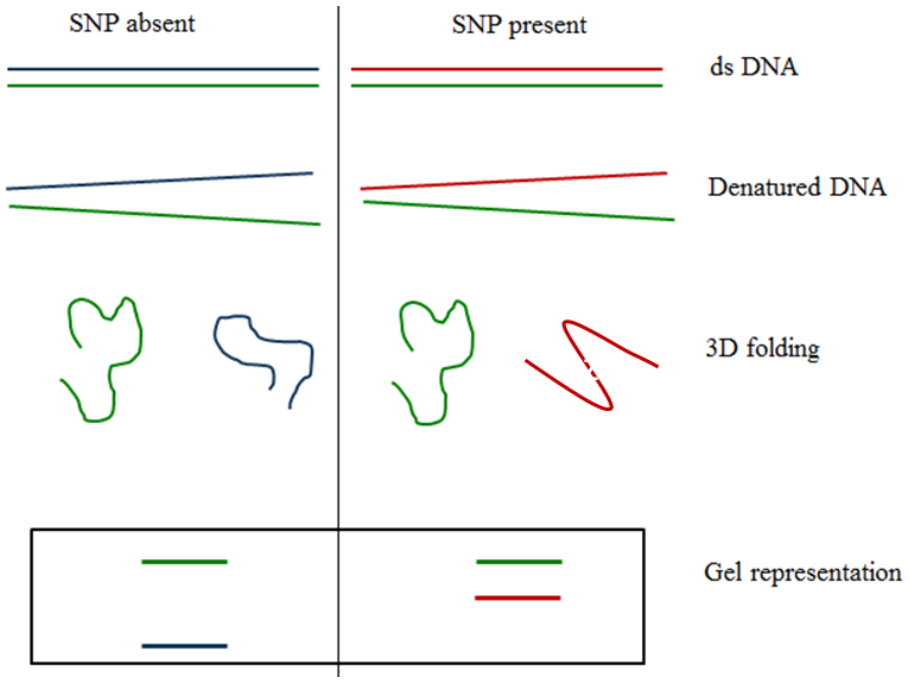


Figure 4.2: Diagrammatic representation of SSCP in a locus with (blue) and without a SNP (red).

Once a SNP has been found in a locus, allelic expression will be analysed using cDNA.

4.2.1 Candidate genes for worker reproduction

A literature search was conducted to find candidate genes that were differentially expressed in queens, reproducing workers and non-reproducing workers (Table 4.1) of different social insect species. In addition, UPF0047 protein yjbQ, long-chain fatty acid transport protein 4 and acyl-CoA synthetase from the AIMS results (Chapter 3) were also used as candidate genes.

Table 4.1: Candidate genes from the literature search. Each gene is differentially expressed between reproducing and non-reproducing insects. Accession numbers are also included.

GENE	REASON	REFERENCE
Major royal jelly protein *(MRJPII)* AF000632	-Differentially expressed in brain tissue of sterile and anarchistic workers.	-Wang *et al.,* 2012
Major royal jelly protein *(MRJP1)* NM_001014429	-Differentially expressed in brain tissue of sterile and anarchistic workers.	- Wang *et al.,* 2012
Vitellogenin	Over-expressed in abdomens of ovary activated workers compared to workers whose ovaries were experimentally inhibited. Over-expressed in *Lasius niger* & *Polistes canadensis* queens. Expression higher in egg laying queens than non-egg laying queens in *Solenpsis invicta*.	-Koywiwattrakul *et al.*, 2005 -Graff *et al.*, 2007 -Sumner *et al.*, 2006
Chymotrypsin DN048380	Up-regulated in *B. terrestris* non-reproducing workers when compared to reproducing workers	-Pereboom *et al.*, 2005
Cytochrome oxidase DN048381	Up-regulated in *B. terrestris* non-reproducing workers when compared to reproducing workers	-Pereboom *et al.*, 2005
Peroxiredoxin DN048382	Up-regulated in *B. terrestris* non-reproducing workers when compared to reproducing workers	-Pereboom *et al.*, 2005

4.2.2 Primer construction

The genomic sequence and mRNA sequence for each gene was aligned to find the exons. Primers were made to span the exons where possible. Intron sequences were not

used because they would not be included in cDNA which would be used for the RT-PCR SSCP. Primers were tested prior to running SSCP. The primers are summarised in Table 4.2.

Table 4.2: List of primers used for SSCP. These primers were used to amplify candidate genes associated with worker reproduction. The expected product size is also included

GENE	PRIMER	SIZE
Major royal jelly protein I	MRJP1.1-F ATCGAAGCCACGAACATTCT	385
	MRJP1.1-R CGGGATACGGGATCAGTAAA	
	MRJP1.2-F GCAATCGTTTGTGGATTGTG	392
	MRJP1.2-R GTCGAAAAATCCGATTGGTG	
	MRJP1.3-F CAATTCGTTGGCCAGCTATC	301
	MRJP1.3-R CAACTCTTCTTCGCCCTGTC	
	MRJP1.4-F GACAGGGCGAAGAAGAGTTG	271
	MRJP1.4-R TGCCATCATACGAAAGGACA	
	MRJP1.5-F TCCTTTCGTATGATGGCAAA	206
	MRJP1.5-R AACTCCGCGATACGTATTAATTG	
Major royal jelly protein II	MRJP2.1-F CACGCATATTGCATTTCCAC	291
	MRJP2.1-R TAATGGCAGCTTGCCTTCTT	
	MRJP2.2-F ATCGTTTACGACGGCAGTTC	215
	MRJP2.2-R TGAAGCTGCGTCAACTGGTA	
	MRJP2.3-F AGCGGTCTCAAAGTGAAGGA	256
	MRJP2.3-R TCCCCGACCTACGATATAACC	
	MRJP2.4-F ACTGCGAGCACAAATACAGC	222
	MRJP2.4-R CGATGACGATCAAAATGTGG	
Yolkless 1	VgR1.1-F TTTTCGTTTAACAGTCGTAATGG	285
	VgR1.1-R TCGGAACATTGGTACTCGTCT	
Yolkless II	VgR1.2-F GCGACGCAAAAGAAGATTGT	160
	VgR1.2-R CATCTTTTACGCAACCATCG	
Peroxiredoxin I	Pox-F TGAAGAAAGCGGAGTTCCAT	172
	Pox-R CCAGCCAGCAGGACATACTT	
	Pox2-F TTCAGAAACCTGCTCCTGCT	163
	Pox2-R GCACGATCAGAAAAGGCAAT	
	Pox3-F TGCACCAATTAAAAAGATTCCT	123
	Pox3-R GGGAATGAAAATAGTTCATATTGAAAA	
Peroxiredoxin IV	Pox1.3-F TATCTTCGTGAACGCCACTG	280
	Pox1.3-R CGCAAGGTCTGTCTAGTCCA	
Peroxiredoxin IV	Pox1.1-F ACGACTTTCGACCCGTTTCT	250
	Pox1.1-R TGATTTCCGTAGGGCAGATA	
	Pox1.2-F CTCCGTACGCACATCTAGCA	325

(continued)	Pox1.2-R AGCTGGTCGCTTGTATCGTT	
Vitellogenin	Vit-F GTGACAAGCGAAGAGACTATTATG	134
	Vit-R CCGTGTTATCTGGCGTGAC	
	Vit2-F TTCTCGAGCCTTACCTGGAA	193
	Vit2-r GGATCGGTCTGCATCAGAAT	
	Vit3-F GATGGACGACGAGGTTTTGT	196
	Vit3-R GGGAGCCATAAAGTCGTTGA	
	Vit4-F GCCACGAAGGGTCAGAATTA	158
	Vit4-R TTG CCC ATG TAG CCA ATG TA	
Lipoprotein receptor	LipR.1-F TCGTTTGCAGGTGATAGTCG	552
	LipR.1-R TCCGTCTGGTTGTCTGTACG	
Chymotrypsin	Chy-F TTGATTCGCGTGAACAAAAA	150
	Chy-R TTGCAAGTTGTTTGGTGCTC	
Cytochrome oxidase	Cyo-F AGGAATCTCTTCAATTCAATTATTGGATCA	165
	Cyo-R ATGGTAATTGCACCGGCTAA	
UPF0047 protein yjbQ	UPF.1-F AGCAGATGCAATTAGACAATCTTT	307
	UPF.1-R CAAAGTCCGACTGAAAATTCG	
long-chain fatty acid transport protein 4	Lcfat.1-F AAAAAGAAATCTCATGGGAACAA	301
	Lcfat.1-R AGCACAATTCCAACCCAGAA	
	Lcfat.2-F TTCATCCTTGATCTGGTTTTCA	337
	Lcfat.2-R CGGAATGATTAATTTCGCAAC	
acyl-CoA synthetase short-chain family member 3	CoA.1-F TTGAACGCCTGTTACAATGC	274
	CoA.1-R CGCCAAATACGACAGAATGA	
	CoA.2-F GACACCATGGATGTTGGCTA	182
	CoA.2-R TCTCCTTTAGTGGGATCTGGAA	

4.2.3. Behavioural assay

The behavioural experiment conducted in section 3.2.3.1 was repeated on four *B. terrestris* colonies.

4.2.4 Amplification and denaturing DNA

Each sample to be amplified was made to a final volume of 100µl and was optimized according to section 2.3.1 by varying annealing time and temperature. The products were visualised according to section 2.2.5 to check if the products were of the expected size.

A denaturing mixture of 1ml formamide and 10µl 1M NaOH was freshly prepared. PCRs were denatured by mixing 4µl of PCR with 7µl of denaturing mixture and incubating in a thermocycler at 95˚C for 7 minutes then immediately chilling on ice for 5 minutes.

4.2.5 Single strand confirmation polymorphism (SSCP)

SSCP was conducted on the Origins tank using GMA (Wide Mini S-2x25) gels (Elchrom/V.H. Bio), which have a resolution of 1bp. The buffer temperature was set to 4˚C and the gel was placed on the catamaran frame and left in the tank for 30 minutes. The tank pump was switched off and 10µl of the denatured PCR was mixed with 2µl of 1x loading buffer (Elchrom) and loaded onto the gel. The tank was then connected to the powerpack, the pump switched on and run at 72V according to fragment size. The running time was 10 hours for products 150-200bp in length, 12hours for products 200-250bp in length, 15 hours for products 250-350bp in length and 17 hours for products 350-450bp in length. Once the run was finished, the gel was stained with SybrGold (Invitrogen) (1:10000 dilution in TAE) in a staining tray on a rocking platform for 40 minutes. The gel was destained with 100ml TAE, again in the dark on the rocking platform for 30 minutes. The gel was visualised under UV light on a transillumiator.

4.3 RESULTS

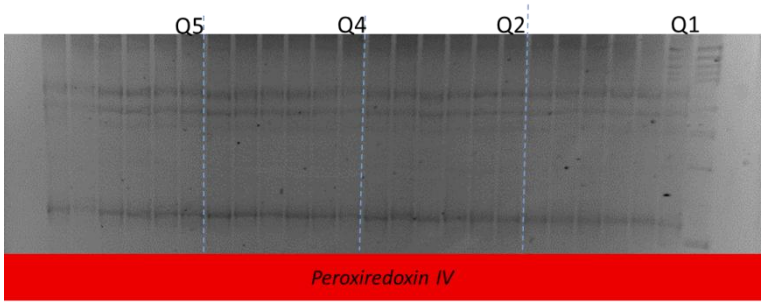
Table 4.3 shows the coverage per gene. The number of genes analysed was given more preference over the percentage of gene covered.

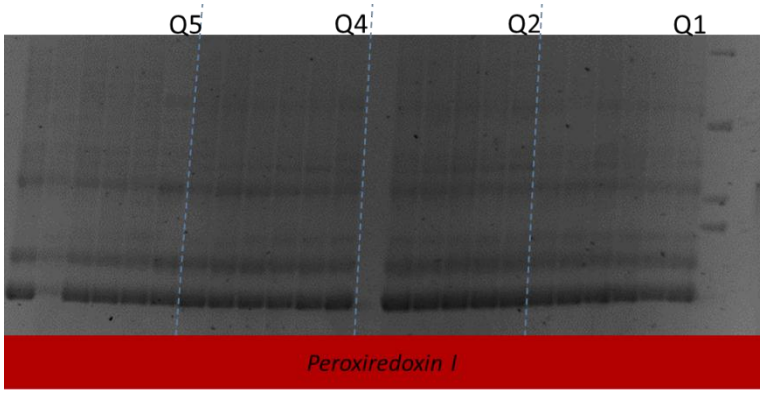
Table 4.3: Exon coverage. The genes in black font were from the literature review, while genes in red were from Chapter 3.

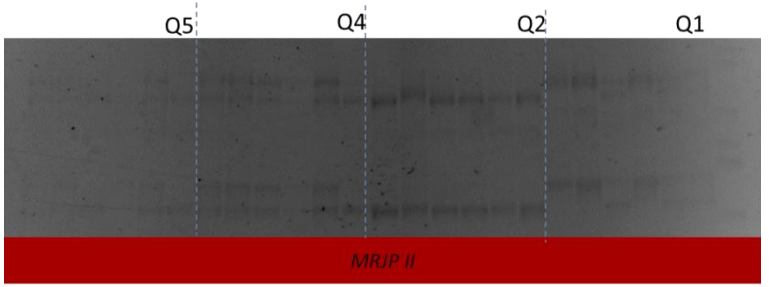
GENE	EXON COVERAGE
Major royal jelly protein I	94.6%
Major royal jelly protein II	38.1%
Yolkless I	5.2%
Yolkless II	3.1%
Peroxiredoxin I	50.1%
Peroxiredoxin IV	83.8%
Cytochrome oxidase	27.4%
Chymotrypsin	34%
Vitellogenin	12.5%
Lipoprotein receptor	54.4%
UPF0047 protein yjbQ	39.9%
long-chain fatty acid transport protein 4	14.8%
acyl-CoA synthetase short-chain family member 3	18.5%

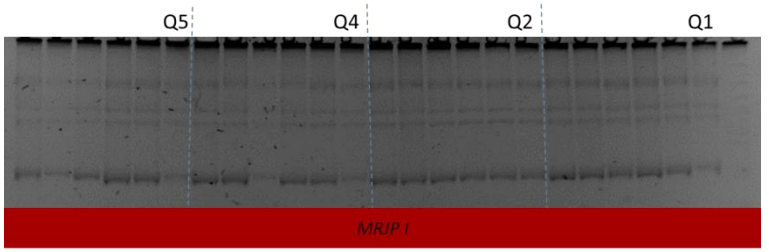
Figures 4.3 and 4.4 show gels from a number of genes for 4 colonies. In each gel, 5 worker samples are followed by the queen from the colony. Variation was not present in any of the sections amplified as all of the bands found in the queen's sample were also present in workers from that colony. If variation had been present, unique bands would be found in the workers which were absent in the queen, which would have been the paternal allele.

Traditionally, SSCP involved the using radio-labelled primers. However, using 'cold' SSCP (without radio-labelled primers) increases sensitivity. In addition, 'cold' SSCP takes less time to run, is inexpensive, safer and more reproducible (Hongyo *et al.*, 1993). As evident in Figure 4.3 and Figure 4.4, results can range from only two strands being produced on a gel to multiple bands. Multiple bands are usually a result of secondary and tertiary structures forming between the nucleotides due to base pairing within the strand. This is also indicative of different alleles being present, as well as multiple confirmations within the locus (Gasser *et al.*, 2006).









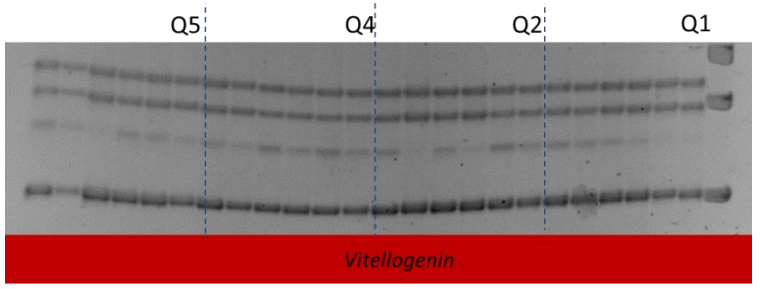
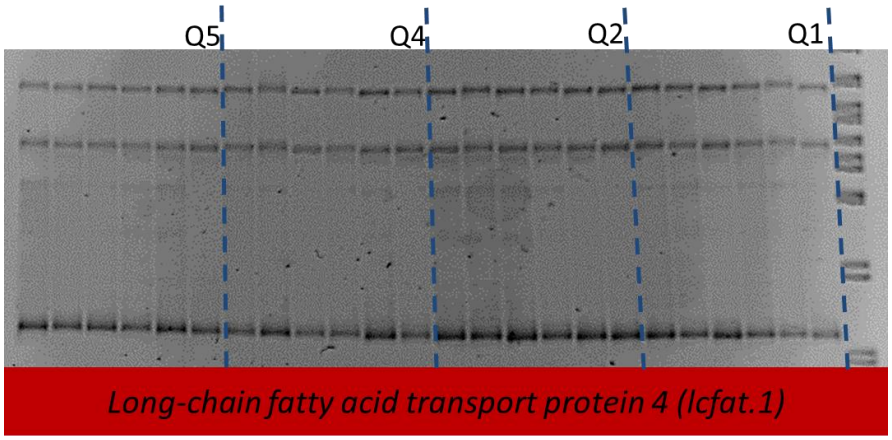
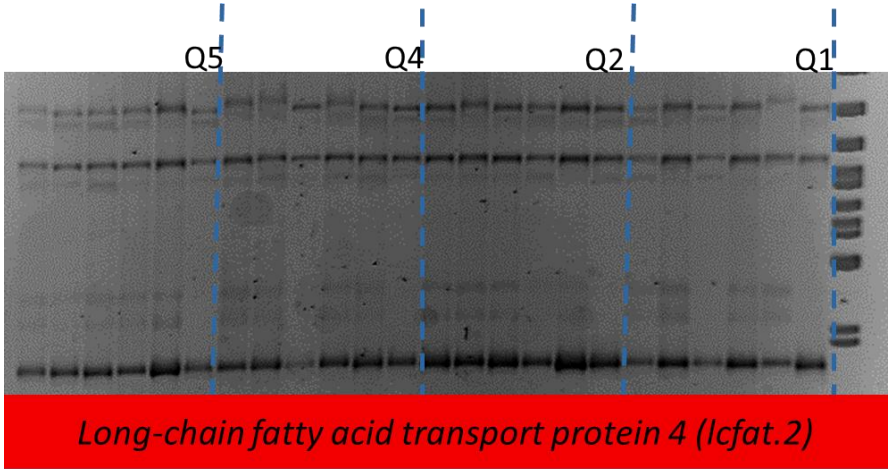
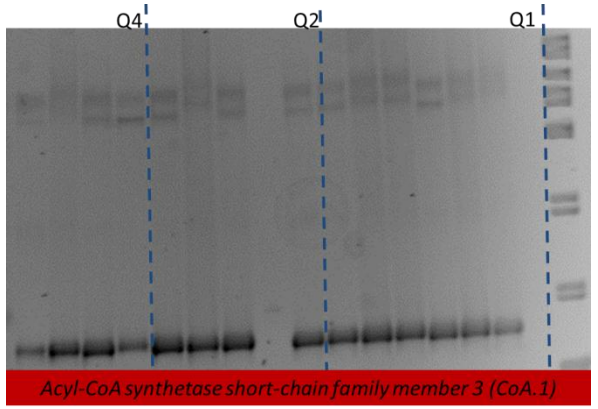


Figure 4.3: SSCP gel results from a single primer set for 5 genes. Each queen is preceded by 5 workers from the same colony. Q1= queen from colony 1, Q2= queen from colony 2, Q4= queen from colony 4 and Q5= queen from colony 5. Variation was not found in any of samples and therefore maternal and paternal loci could not be distinguished.







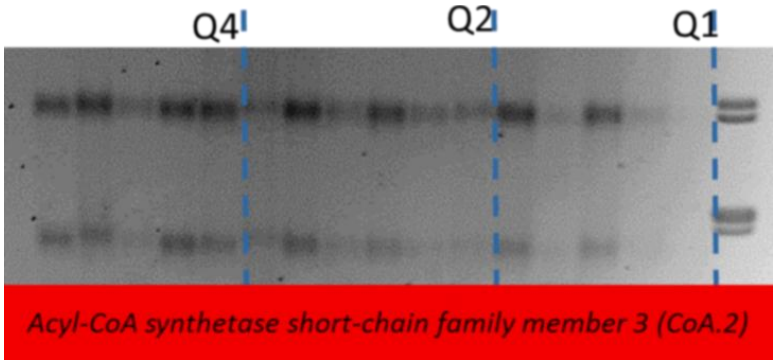


Figure 4.4: SSCP gel results from 2 of the genes from AIMS. Each queen is preceded by 5 workers from the same colony. Q1= queen from colony 1, Q2= queen from colony 2, Q4= queen from colony 4 and Q5= queen from colony 5. Variation was not found in any of samples and therefore maternal and paternal loci could not be distinguished.

4.4 Discussion

4.4.1 Summary of results

The aim of this chapter was to test Haig's conflict theory by assessing parent-of-origin allele specific expression in genes associated with reproduction in queen-less *Bombus terrestris* workers. The prediction, that genes which promote worker reproduction will have an imprinted paternal allele while the maternal copy will be expressed, could not be tested because the alleles could not be distinguished.

Sections of the candidate genes were amplified successfully and tested for SNPs using SSCP. Unfortunately, the samples did not exhibit variation, therefore making the identification of maternal and paternally inherited alleles impossible. The primers used targeted exons in this study because methylation in insects was found in the coding regions of *Apis mellifera* and *Acyrthosiphon pisum*. Rarely has methylation been found in non-coding intron regions (Lyko *et al.*, 2010). It is also generally believed that any variation, which was a key aspect of this experiment, would be found in the exons. In addition, exons were amplified because the ultimate part of the experiment was to look at expression differences in the alleles of the candidate genes.

Unfortunately, SNPs were not identified in this experiment and the final step involving cDNA could not be conducted. The theory behind RT-SSCP PCR is identical SSCP except for the use of cDNA instead of DNA, therefore making the allele that is not expressed (imprinted) absent in the gel (Figure 4.5).

Figure 4.5: Hypothetical RT-PCR SSCP results. If a locus is paternally imprinted, only the locus which is not imprinted will be present on the gel. (Q= queen, M= male & W= worker)

It has been documented that point mutations in GC rich areas are more difficult to find (Gasser *et al.*, 2006) which may have influenced the ability to detect SNPs in *B. terrestris.* Another reason for SNPs not being identified might be because the bees used were obtained commercially, usually results in low genetic variation compared to bees in the wild. Initially, approximately 200 queens were captured in early spring with the intention of rearing colonies for this experiment. They were collected from the Botanical Gardens in Leicester, Keele University and Calderstone's Park in Liverpool to ensure diverse populations. Regrettably, none of the colonies survived for more than a few weeks.

4.4.2 Implications and further work

The candidate genes chosen in this study to test for imprinting were all associated with worker reproduction. The plasticity caused by the differentially expressed genes in reproducing and non-reproducing members of eusocial insect species may be due to methylation, and therefore be ideal to test for imprinting. Supporting this is that three of the candidate genes used in Chapter 5 were found to be methylated in Chapter 3. A recent study found 1292 genes to be differentially expressed in non-reproductive and reproductive workers in queen-less *A. mellifera* colonies (Cardoen *et al.*, 2011). The

candidate genes used in this experiment were ones that were found to be differentially expressed in queens, reproducing workers and non-reproducing workers across various social insect societies. *MRJP* (major royal jelly protein) genes are found in both *A. mellifera* and *Nasonia vitripennis* (Nasonia working group, 2010) and this set of genes is associated with caste development and social behaviours. *MRJP* in *A. mellifera* is also associated with reproductive maturity. It has found to be usually expressed highly in workers, but sometimes in queens, suggesting that this gene has diverse issue dependent expression patters (Elango *et al.*, 2009).

Vitellogenin, vitellogenin receptors and lipoprotein receptors were used as candidate genes in this study because of their importance in insect reproduction. Vitellogenin is synthesised in the fat body and is used by oocytes by a process of receptor-mediated endocytosis. Vitellogenin has been suggested to be involved with the transport of carbohydrates, lipids, vitamins, hormones and metals, as well as its traditional role in oocyte maturation (Guidugli-Lazzarini *et al.*, 2008). Coupled with juvenile hormone, a decrease in vitellogenin expression can lead to *A. mellifera* workers to become foragers (Nelson *et al.*, 2007). Vitellogenin receptors are also important in bee biology as they are directly involved with vitellogenin expression. These receptors are usually low density lipoprotein receptors (LDLRs) which are only expressed in ovaries (Ciudad *et al.*, 2006). Lipoprotiens are another protein which is involved in moving lipids from the fat body to where they are needed- the developing oocytes. Again, like vitellogenin receptors, lipoprotein receptors are also LDLRs. The lipoprotien receptor in non-reproducing *A. mellifera* workers have been found to upregulated compared to reproducing workers (Cardeon *et al.*, 2011).

Bombus impatiens queens can sometimes mate twice (Cnaani *et al.*, 2002). In addition, while 50% of *B. terrestris* workers in queen-right colonies had mature oocytes post-competition point, only 9% of *B. impatiens* workers had mature oocytes at the same phase. It would be interesting to repeat the experiment in this chapter to see how polyandry would affect allelic expression. Here the prediction would be that the patrigenes would be as selfish as the matrigenes in workers from queen-less colonies and no imprinting would occur (Queller, 2003).

Using SSCP and RT-PCR SSCP is one of many ways to find imprinted genes. Identifying imprinted genes can be challenging because monoallelic expression may only occur in a specific isoform, tissue or stage of development. Thus, not finding evidence of imprinting at a particular part of a gene in a particular tissue or life stage does not necessarily mean that imprinting is not present. In addition, imprinted genes are rare, with imprinting in the human genome has been demonstrated to occur approximately 1% of the time (Luedi *et al*., 2007). Recent advances in DNA sequencing technologies can be utilised to find novel imprinted genes. However, this is slightly more complicated because the *B. terrestris* genome is only partially annotated.

Chapter 5 : General Discussion

5.1 A summary of the results

The main objective of this thesis was to use *Bombus terrestris* as an insect model in the validation of Haig's conflict theory. The first aim of this thesis, addressed in Chapter 2, was to establish the presence of a methylation system in *B. terrestris*. After confirming the presence of methytransferase genes, Chapter 3 assessed the presence of methylated genes in different caste and life stages. Finally, once this evidence for a platform for establishing imprints was shown to be present in *B. terrestris,* Haig's conflict theory was tested using worker reproduction in queen-less colonies in Chapter 4. The following sections revisit the main findings and then discuss the implications and further application of this thesis.

5.1.1 The *Bombus terrestris* methylation system

The first experiment used a combination of PCR and BAC library screening to show that a functional vertebrate –like methylation system exists in *B. terrestris.* The methylation system in *B. terrestris* is composed of *Dnmt1a*, *Dnmt1b*, *Dnmt2* and *Dnmt3,* as well as a MDB protein. The *B. terrestris* methylation system reflects the methylation systems found in *B. impatiens* and *Apis mellifera*, in that all three methyltransferases are present, with two orthologues of *Dnmt1*.

5.1.2 Methylated genes in *Bombus terrestris*

The next study assessed the methylomes of various larval stages, pupae, callow workers, queens, workers and queen-less reproducing workers as well as reproducing workers in a queen-right colony. Differential methylation profiles were produced when the DNA was subjected to methylation sensitive and insensitive restriction

endonucleases. When sequenced, these differentially methylated loci were revealed to code for proteins with diverse functions- binding, cellular components, enzymes, membrane or structural components and transport.

5.1.3 Testing Haig's conflict theory

Finally, this thesis set out to test one facet of the genomic conflict theory- that in genes involved with reproduction in workers of queen-less *B. terrestris* colonies, the paternally derived allele would be imprinted, while the maternally derived allele would be expressed. RT-PCR SSCP was to be used to check allele expression once alleles were genotyped using SSCP. Unfortunately, variation was not present in any of the loci that were tested and thus the maternally derived allele could not be distinguished from the paternally derived allele. Imprinting is the parent-of-origin dependent silencing of alleles and examining expression differences in the alleles without knowledge of which parent each allele was inherited from would be futile.

5.2 Discussion of results

DNA methylation systems in insects have also provided some surprises. For example, *Bombyx mori* and *Tribolium castaneum* both have a similar compliment of methyltransferases- one orthologue of *Dnmt1* and *Dnmt2* and no *Dnmt3*. It is universally agreed that both *Dnmt1* and *Dnmt3* are necessary to create and maintain cytosine methylation. This idea is mirrored in *T. castaneum*, which does not show any methylation. However, less than 0.2% cytosines have been found to be methylated in *B. mori*, usually in gene bodies that code for small RNAs (Lyko & Maleszeka, 2011). This

questions how relevant methylation is for the functioning of genes, as well as how the lack of a *de novo* methyltransferase can still lead to methylation.

It can be argued that finding methylation in *B. terrestris* just fills another blank in the evolutionary tree of insect methylation systems. However, *B. terrestris* provides a unique biological and behavioural perspective when compared to other hymenopteran insects. For example, while *B. terrestris* and *A. mellifera* are both hymenopteran species, *B. terrestris* queens are singly mated while *A. mellifera* is a polyandrous species. *Acyrthosiphon pisum* like *B. terrestris* is eusocial, but has a more complex social structure that can interfere with the fine tuning of epigenetic studies. Thus, *B. terrestris* methylation, coupled with their distinctive biology offers us a unique opportunity to explore and understand the role epigenetics plays in eusociality.

Bombus terrestris workers may not lucidly behave in their kin-selection conflicts. Instead, the queen uses pheromones to stifle egg laying by workers until the queen's interests are served in the form of laying males and new queens (Bourke & Ratnieks, 2001). The egg laying behaviour of the workers and conflict has been suggested to be a by-product of the queen's decrease in dominance. Cooperation and conflict go hand-in-hand in social insect colonies. Most workers forego their own reproduction and invest time and resources to take care of the queen's offspring. However, due to reproductive plasticity some workers produce parthenogenic males. The extent to which this plasticity is motivated by epigenetics requires further analysis. There is an uncertainty regarding the optimal strategy for imprinting in social insects because the final fate of the individual (haploid male, diploid worker or queen) is unknown. This can cause

problems when hypothesising where the imprint should occur during gametogenesis. Thus, it has been suggested that the predictions for genomic imprinting in social insect systems may be more intricate than we believe (Kronaur, 2008).

Queens and workers share many of the genes that are expressed (Hoffman & Goodisman, 2007). Therefore, not only are the genes involved in the development of larvae, but also in behaviours such as killing of male larvae, feeding of larvae and other developmental genes. It cannot be predicted how genes affecting worker and queen development and behaviour can affect male development and behaviour. One way to explore this would be to test differences in epigenetic patterns between males produced by workers and queens (Kronaur, 2008).

5.3 Applications

Finding imprinted genes in insects can be an asset to understanding mammalian epigenetics. For example, ageing is proposed to be a result of a step-by-step silencing of housekeeping genes (Krȍll, 2007). *Drosophila melanogaster* have been shown to live longer when Hsp22 and Hsp70 expression is increased via the application of histone deacetylase inhibitors (Zhao *et al.*, 2005). These histone deacetylase inhibitors are now being used in epigenetic cancer therapies (Krȍll, 2007). A slightly different approach to housekeeping genes influencing aging is shown by *A. mellifera* queens who can live for longer, compared to sister workers although they are genetically identical (Chittka & Chittka, 2010). The combined effect of environment and an individual's genetics plays a major role in the study of aging (Mugatroyd *et al.*, 2010) and this area is possibly one

way of applying the apparent methylation-based phenotypic plasticity in insects (Kurcharski *et al.*, 2008 and Robinson *et al.*, 2011).

Reproducing workers and non-reproducing workers in the monoandrous *B. terrestris* species are similar but are able to generate different methylation profiles. In addition, methylation profiles differed throughout the development stages. This could possibly be applied to the study of methylation differences in monozygous twins, who have also shown evidence for age-related methylation differences (Fraga *et al.*, 2005).

From my point of view as an evolutionary biologist, the impact of finding imprinted genes would have interesting ramifications for the method we use to study social evolution in social insect societies. Currently, relatedness is a key component in the analysis of social evolution. For example, bumblebee workers have a relatedness value of 0.75 and numerous predictions have been made regarding the role this relatedness will play in behaviours in the colony such as worker reproduction and sex allocation. However, if instead of considering r=075 we considered 1 (patrigene) and 0.5 (matrigene), a restructuring of the underlying theory will be crucial.

5.4 Further work using next generation technology

Next generation sequencing technologies provide a range of methods for the analysis of methylation and gene expression. Methylomes can be analysed using methyl-seq, reduced representation bisulphite sequencing (RRBS), methylated DNA immunoprecipitation sequencing (MeDIP-seq) and finally, methylated DNA binding

domain sequencing (MBD-seq), producing methylomes with a resolution of a single base in humans and *A. thaliana* (Lister *et al.*, 2009). MethylC-seq is a process where bisulphite treated DNA is shotgun sequenced, while RRBS involves *MspI* restriction digestion (Harris *et al.*, 2010). The initial step in both MeDIP and MBD-seq is the enrichment of methylated regions, by the use of immunoprecipitation in the former and MBD binding domain in the latter before being sequenced. These sequencing methods are being increasingly used due to their cost effectiveness and the shear amount of data produced.

Harris *et al.* (2010) used a series of next generation sequencing methods to find methylation based monoalleleic expression. This pipeline is quite similar to the methods used in this thesis although it is on a genome-wide scale. MRE-seq (methyl-sensitive restriction enzyme) and MeDip can be used in conjunction, which is equivalent to the AIMS work I conducted in Chapter 4. In essence, MeDIP, as explained above, enriches methylated cytosine via an anti-methylcytosine antibody, while MRE-seq identifies unmethylated cytosines by a series of methylation sensitive restriction digestions using *HpaII, Hin61* and *Acil* (Harris *et al.*, 2010). A locus is found in both the MRE-seq and MeDIp results would be hemi-methylated, which is a hallmark of genomic imprinting. The next step is to find loci that are both hemi-methylated and heterozygous. This step is equivalent to the SSCP conducted in Chapter 5. Finally, RNA-seq would be used to analyse expression in the selected loci, which would functionally be similar to RT-PCR SSCP. This approach has an advantage over the AIMS/SSCP/RE—PCR SSCP approach in that the heterozygosity of the individual is assessed at a whole-genome level and not limited to candidate genes. Thus, while the experiment in Chapter 5 could not be completed due to a lack of variation in the candidate gene exons I investigated,

the MRE/MeDip/RNA-seq approach would yield all the monoallelically expressed loci found in the entire genome.

In addition, profiling microRNA (miRNA) within the trascriptome can also help further our knowledge of *B. terrestris* phenotypic plasticity. MiRNA are highly conserved sections of non-coding RNA 18-24 nucleotides in length which regulate gene expression. A recent study used deep sequencing of miRNAs in *A. pisum* (Legeali *et al.*, 2010) to assess phenotypic plasticity. Results indicated differential expression of miRNAs in the endocrine system in aphid morphs who reproduced sexually or asexually (Legeali *et al.*, 2010). It would be fascinating to explore the miRNA profiles of *B. terrestris* queens, noon-reproducing workers and reproducing workers in both queenless and queen-right colonies.

5.5 The evolutionary implications of methylation-driven phenotypic plasticity via genomic imprinting

Gene expression is a result of adaptations to information received from both the organism's body and its external environment throughout its entire life. These expression reactions are not only a reply to current events, but to also to past environments. This information is applied through generations- somatic cells retain imprints while germ cells are do not. Genomic imprinting illustrates that information about the parent of origin is transmitted epigenetically from one generation to another (Haig, 2000). The study of the role epigenetics plays on various life history traits will give us a fascinating insight into the cooperation and conflict within social insect

systems, as well as providing a means of testing predictions based on mammalian behaviours.

Phenotypic plasticity, while an important catalyst for speciation, has limitations such as lag time, developmental range and epiphenotype. The production and maintenance, acquisition of environmental cues, pleiotropy and epistasis are a few of the costs incurred by exhibiting phenotypic plasticity (Thibert-Plante & Hendry, 2010; van Burskirk & Steiner, 2009). In addition, phenotypic plasticity can be maladaptive. A high cost of plasticity, strong selection pressures, low dispersal rates and the presence of few loci controlling the non-plastic component of a trait can lead to phenotypic plasticity not evolving in a population. Loci number affects plasticity because decreases in loci number cause allelic effects to increase, thereby resulting in stronger selection on each loci. These conditions however, favour genetic divergence (Sultan & Spencer, 2002; Gourbiere, 2004).

As discussed, epigenetic mechanisms cause changes in gene expression and therefore the phenotype of the next generation without altering the genotype (Finnegan, 2002). Studies conducted on a wide variety of biological systems, such as yeast (Levy & Siegel, 208), plants (Salmon et al., 2008; Keyte et al., 2006), animal development (Ruden *et al*., 2005) animal behaviour (Crews 2008) and humans (Whitelaw & Whitelaw, 2006) have illustrated that the epigenetic changes caused by hybridisation or environmental stress is heritable. Therefore, epigenetics needs to be considered when analysing adaptations to novel environments. If the role of epigenetic variation and epigenetic inheritance in the natural environment is to be understood, a detailed view of

the importance of epigenetics compared to variation in DNA sequences in ecology and the evolution of populations is needed. It has been suggested that an assimilation of molecular epigeneticists, experimental evolutionary ecologists and theoretical population and quantitative geneticists is needed in order to distinguish between epigenetics and DNA sequence-based effects on the evolution of phenotypic traits (Richards *et al*., 2010).

The impact of epigenetics on the current understanding of evolution is important. Methyltransferases are less stable than the enzymes which govern DNA sequences (Reinder *et al*., 2009) which can result in more mutations. The extra mutations have been suggested to cause one of two inheritance systems which influence phenotypic variation (Richards *et al*, 2010). The first inheritance system is epigenetic, which is fast yet unstable. The second is slower but stable, due to DNA sequence-based inheritance (Bossdorf *et al*., 2008) Such an approach would help understand the persistence of novel advantageous phenotypes, which have to overcome selection while still found at a low frequency throughout a population. Such an approach is concurrent with West-Eberhard (2008), which suggested that genes stabilise phenotypic plasticity caused by epigenetics; therefore genes themselves do not initiate evolution.

Centuries ago, Lamarck suggested that in addition to traits acquired during an individual's life being passed to offspring, the cumulative effects of the use or disuse of these traits can effect morphological changes and result in speciation (Haig, 2011). Weismann (1891) however, disagreed with the heritability of acquired characteristics and supported the singular role of natural selection. Weismann did agree that the

environment can influence germ-lines but did not agree that changes in the germ-plasm could be inherited (Weismann, 1891). However, the accepted school of thought is that natural selection plays a moderate part in causing evolutionary change (Bowler, 1983). While Lamarck was discredited, current understanding of epigenetics, phenotypic plasticity and inheritance may support Lamarck's original hypothesis. If Lamarck was indeed correct, evolutionary theory needs to be substantially modified.

5.6 Conclusion

This thesis has demonstrated that *Bombus terrestris* has a functional methylation system capable of supporting phenotypic plasticity. The studies conducted in this thesis have also demonstrated that genes are methylated differently in *B. terrestris* life stages and caste. In addition, manipulating the hierarchy of a bumblebee colony can result in changes to the division of labour, which is also reflected in the methylation profiles of the workers. The activation of ovaries by which worker reproduction is enabled is not only an exhibition of phenotypic plasticity, but also a means of further influencing the transfer of resources from a mother to her young. These genes which promote worker reproduction were proposed to be governed by Haig's conflict theory but could not be verified. As the framework for genomic imprinting, the presence of a functional methylation system, has been erected in this thesis, future work can directly focus on assessing the validity of Haig's conflict theory.

Appendix: AIMS gels

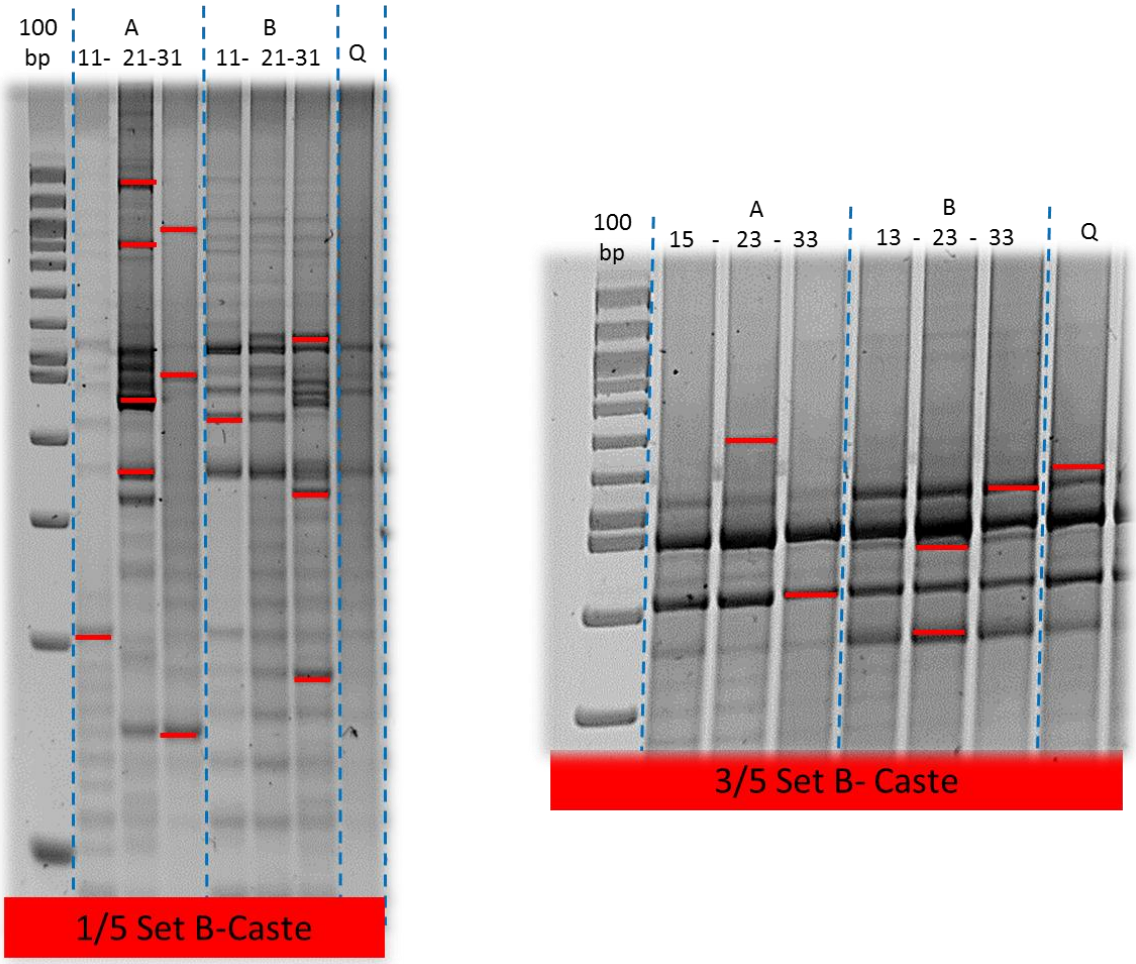


Figure A1: 2 AIMS gels for replicate 1 and 3, using primer set B.

A=10 day old reproducing workers, B= 10 day old non-reproducing workers and Q= queen

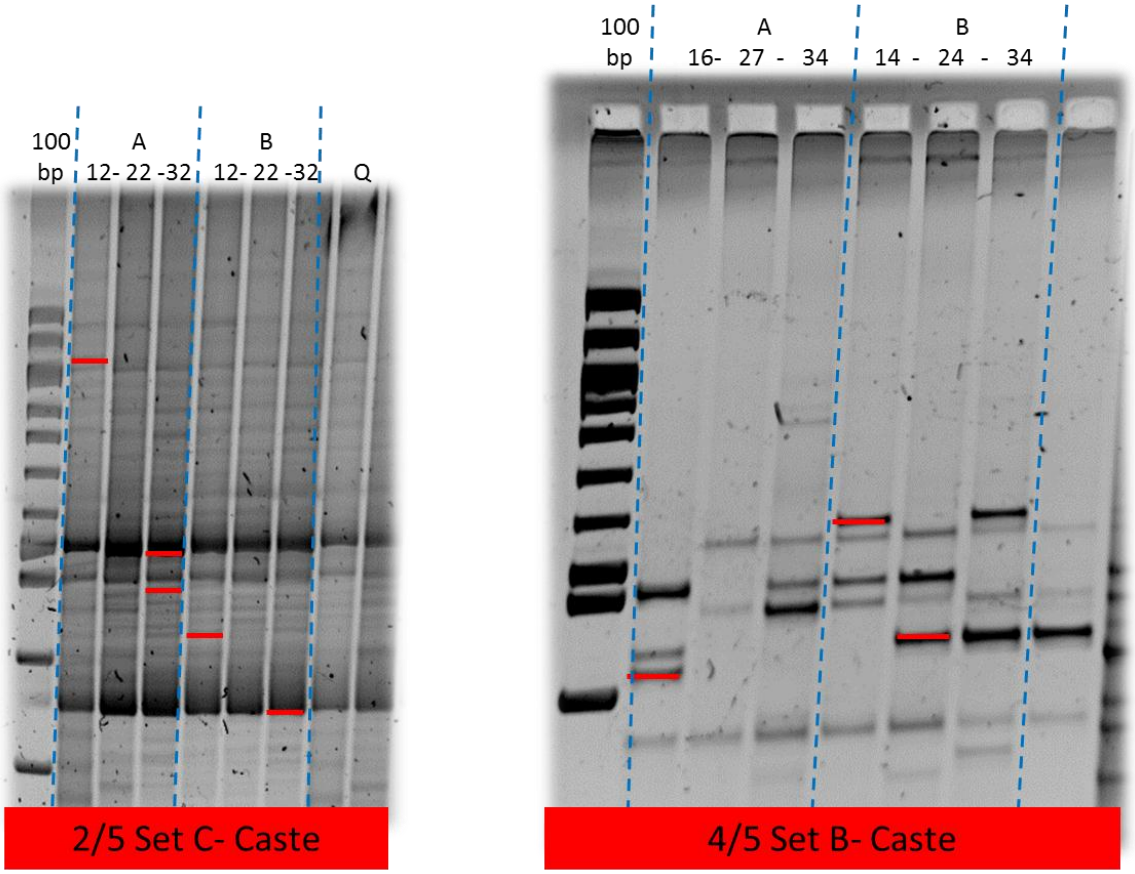


Figure A2: 2 AIMS gels for replicate 2 and 4, using primer set C & B.

A=10 day old reproducing workers, B= 10 day old non-reproducing workers and Q= queen

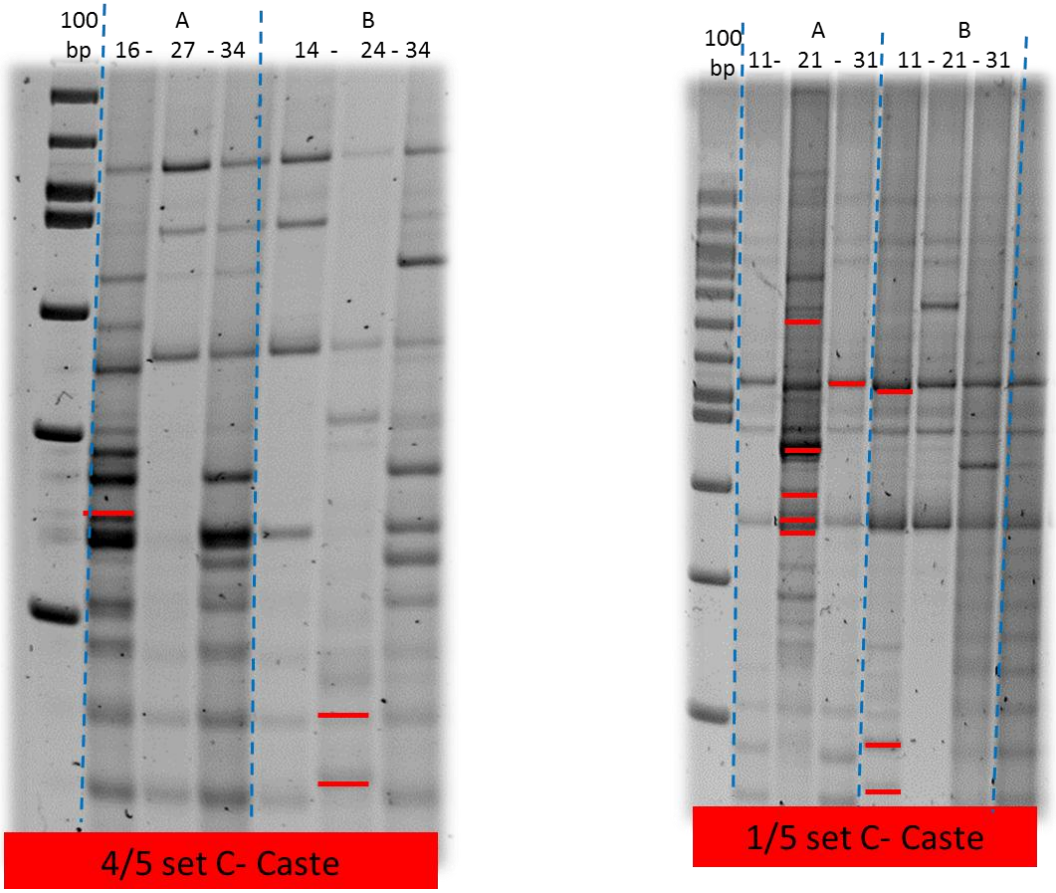


Figure A3: 2 AIMS gels for replicate 4 and 1, using primer set C.

A=10 day old reproducing workers, B= 10 day old non-reproducing workers and Q= queens

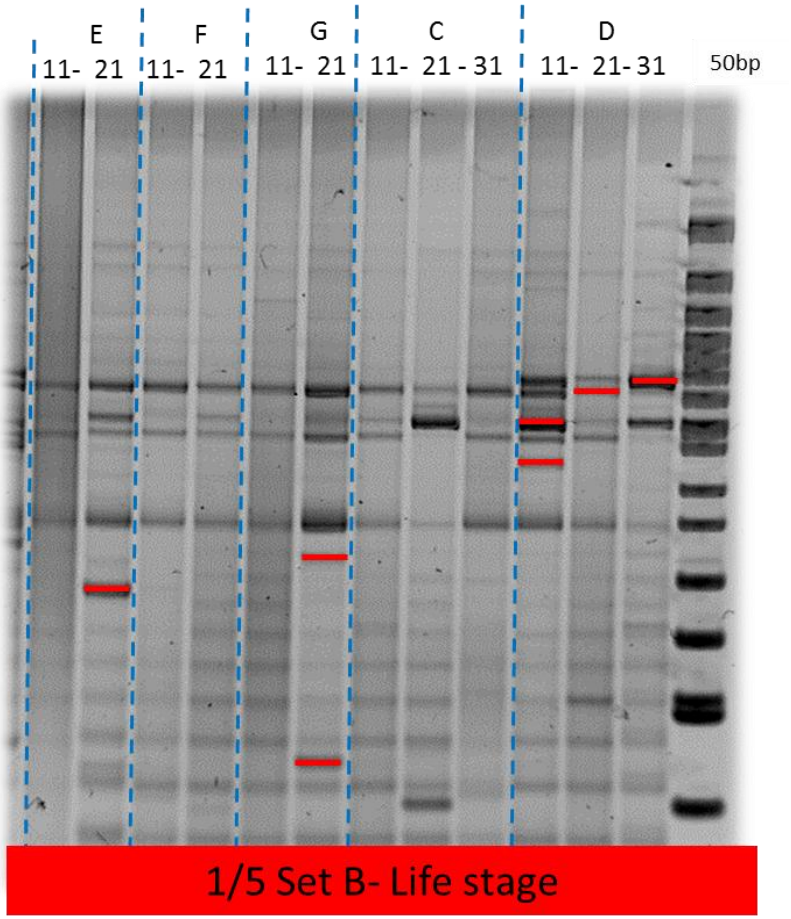


Figure A4: AIMS gel for replicate 1 using primer set B

C= callow workers, D= pupae, E=L4 larvae, F= L3 larvae and G= L2 larvae

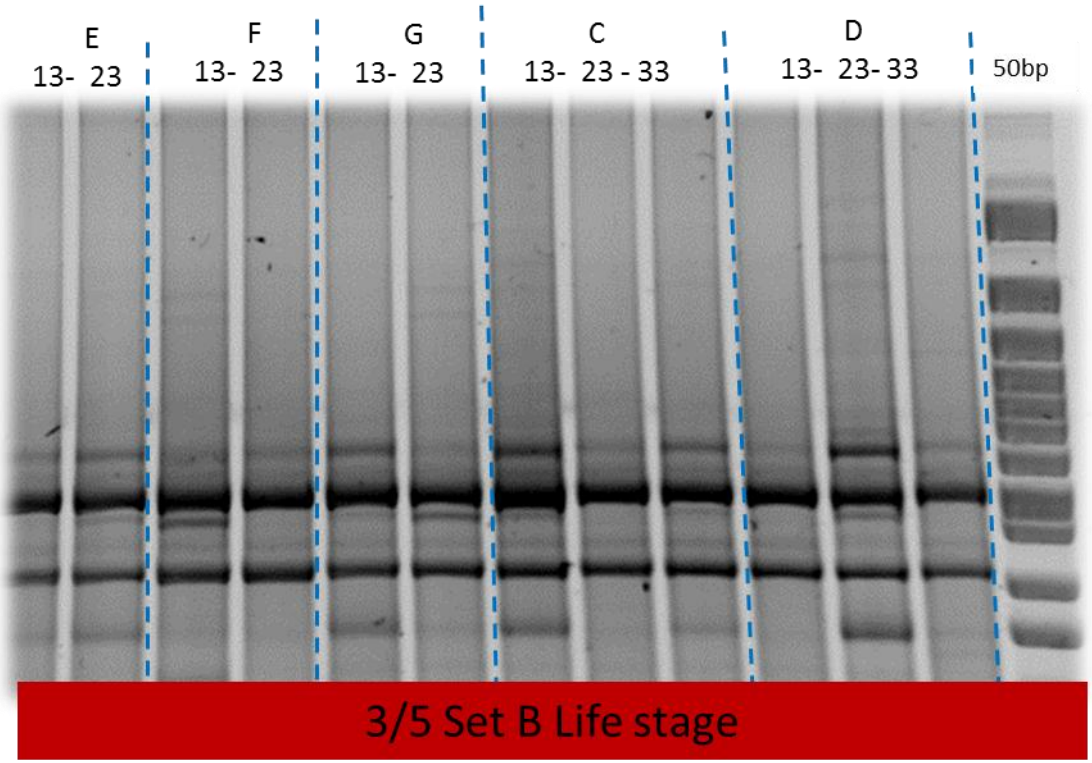


Figure A5: AIMS gel for replicate 3 using primer set B

C= callow workers, D= pupae, E=L4 larvae, F= L3 larvae and G= L2 larvae

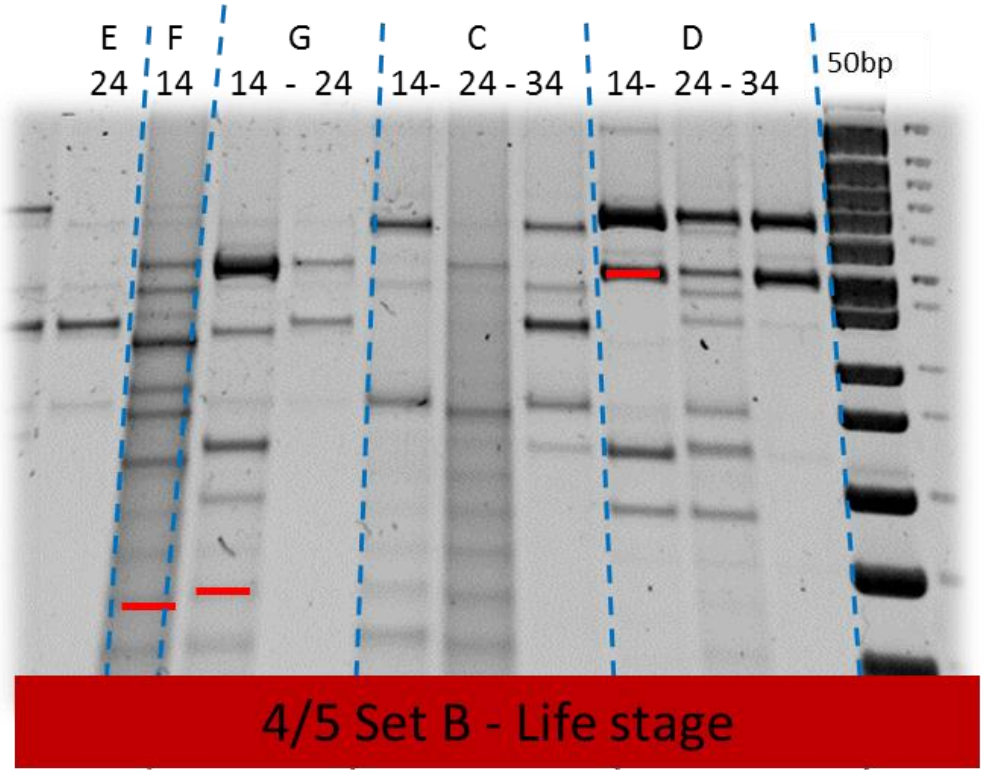


Figure A6: AIMS gel for replicate43 using primer set B.

C= callow workers, D= pupae, E=L4 larvae, F= L3 larvae and G= L2 larvae

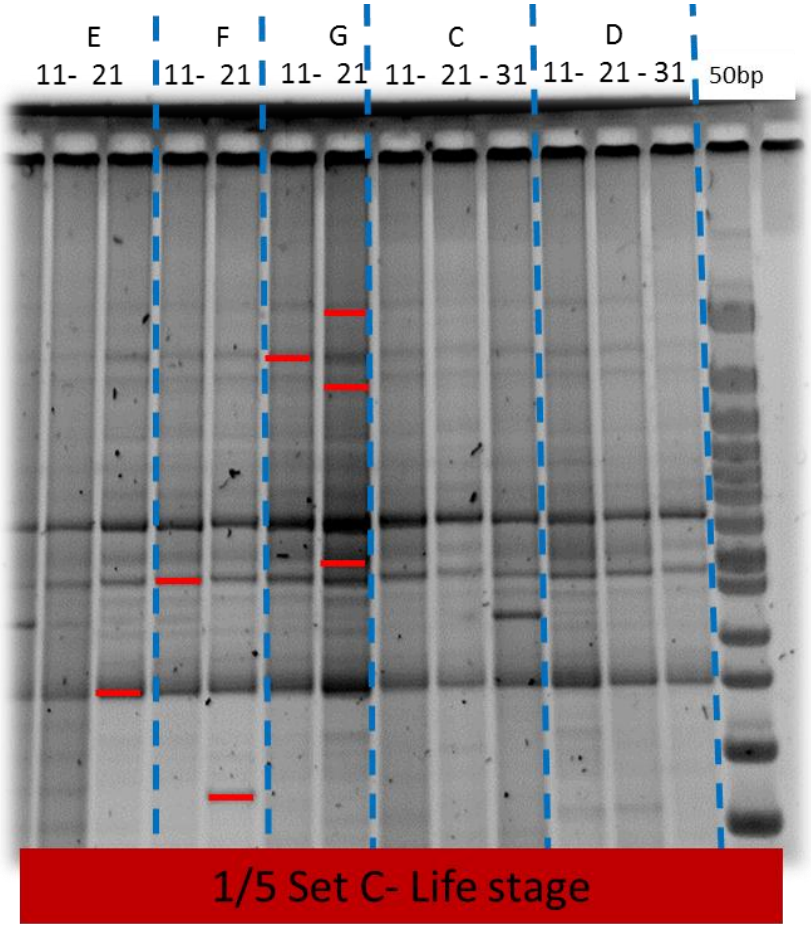


Figure A7: AIMS gel for replicate 1 using primer set C.

C= callow workers, D= pupae, E=L4 larvae, F= L3 larvae and G= L2 larvae

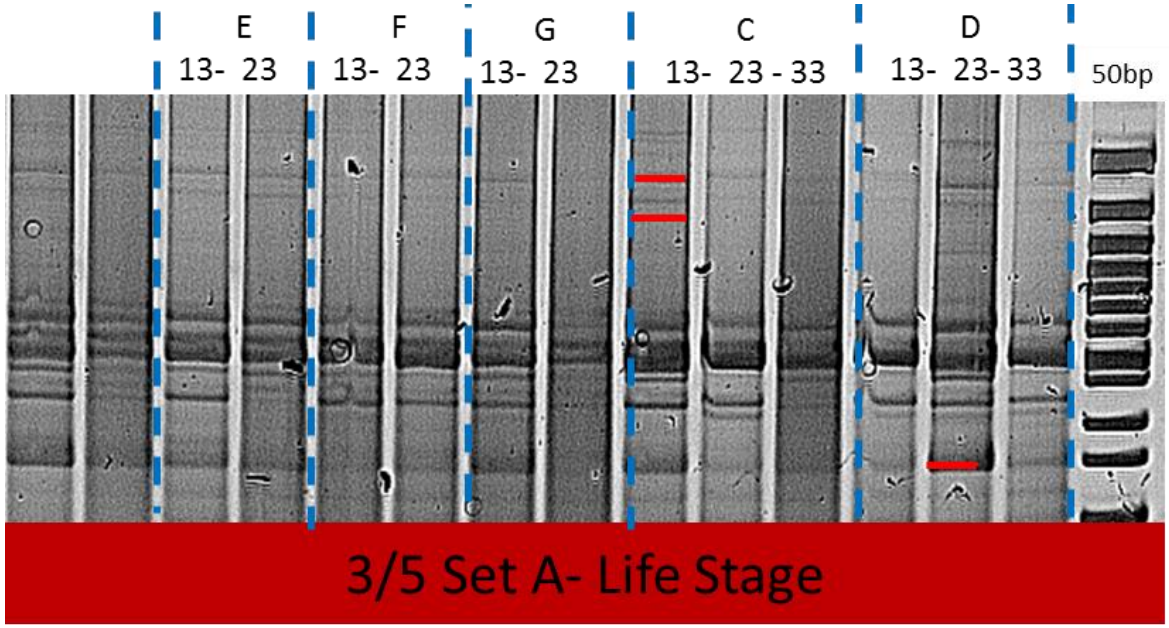


Figure A8: AIMS gel for replicate 3 using primer set A.

C= callow workers, D= pupae, E=L4 larvae, F= L3 larvae and G= L2 larvae

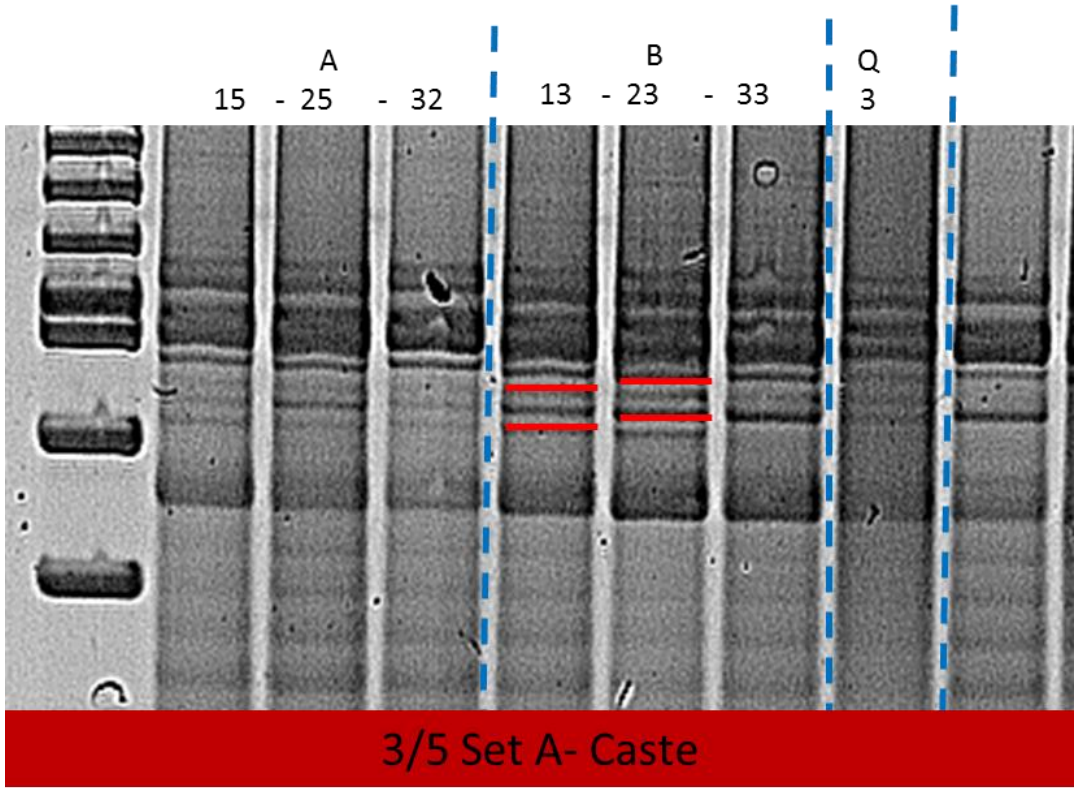
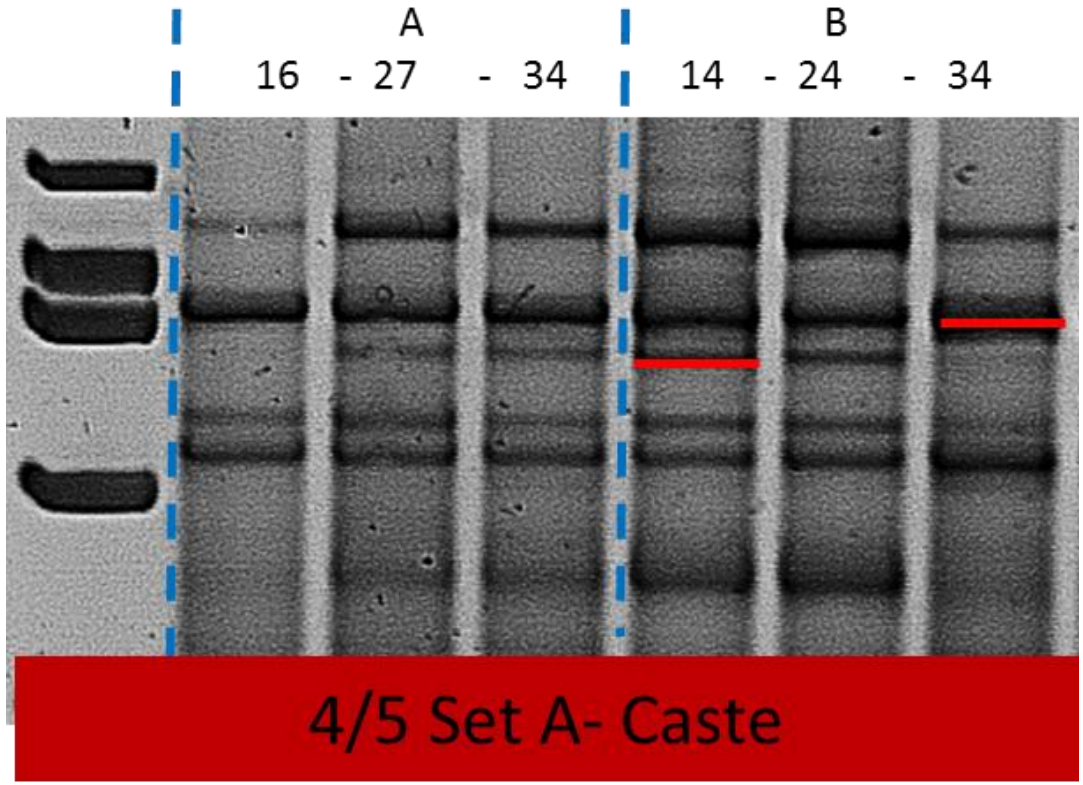


Figure A9: AIMS gel for replicate 3 using primer set C.

A=10 day old reproducing workers, B= 10 day old non-reproducing workers and Q= queen



#

Figure A10: AIMS gels for replicate 4, using primer set C.

A=10 day old reproducing workers & B= 10 day old non-reproducing workers

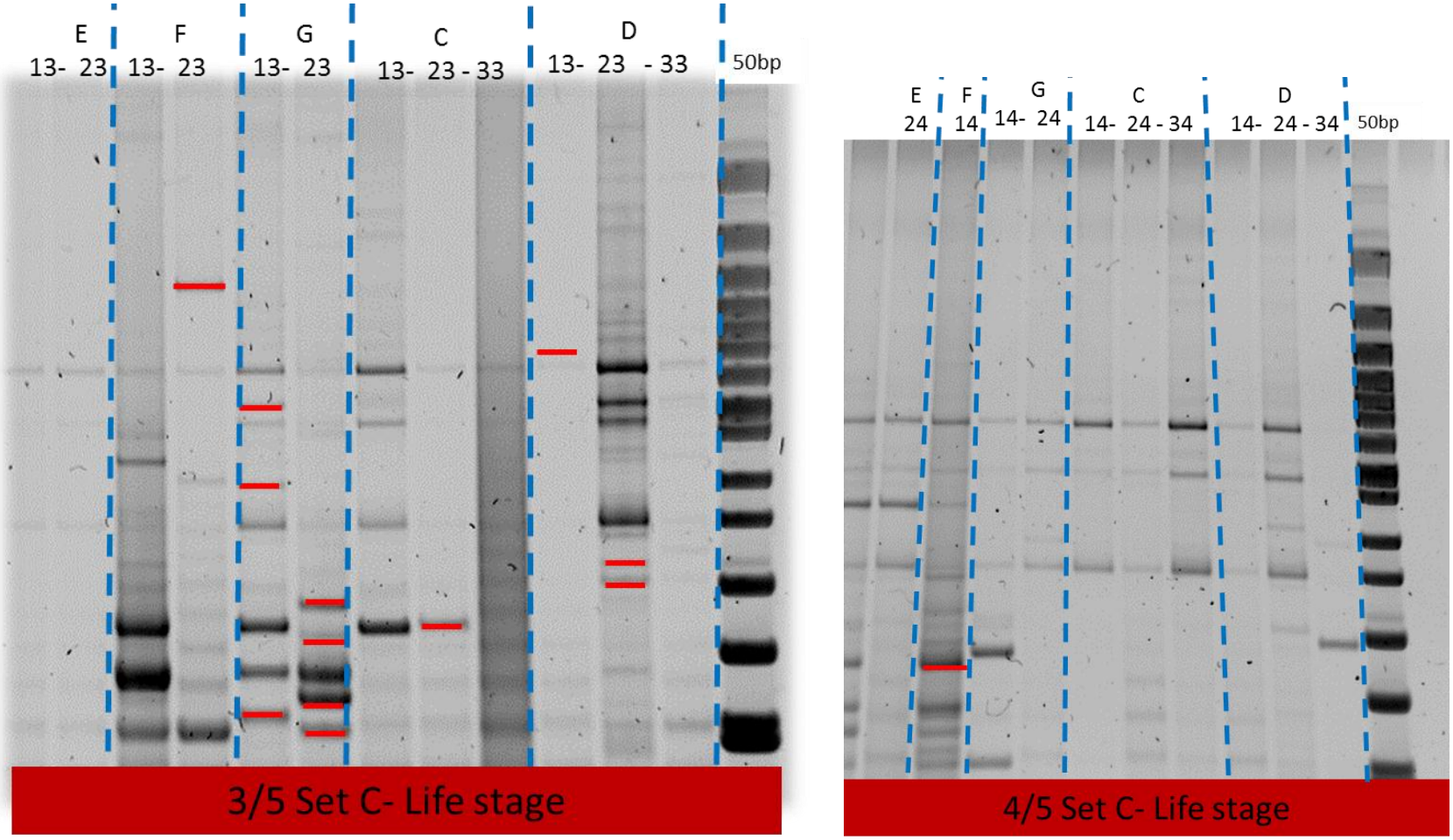


Figure 4.11: 2 AIMS gels for replicate 3 and 4, using primer set C.

C= callow workers, D= pupae, E=L4 larvae, F= L3 larvae and G= L2 larvae

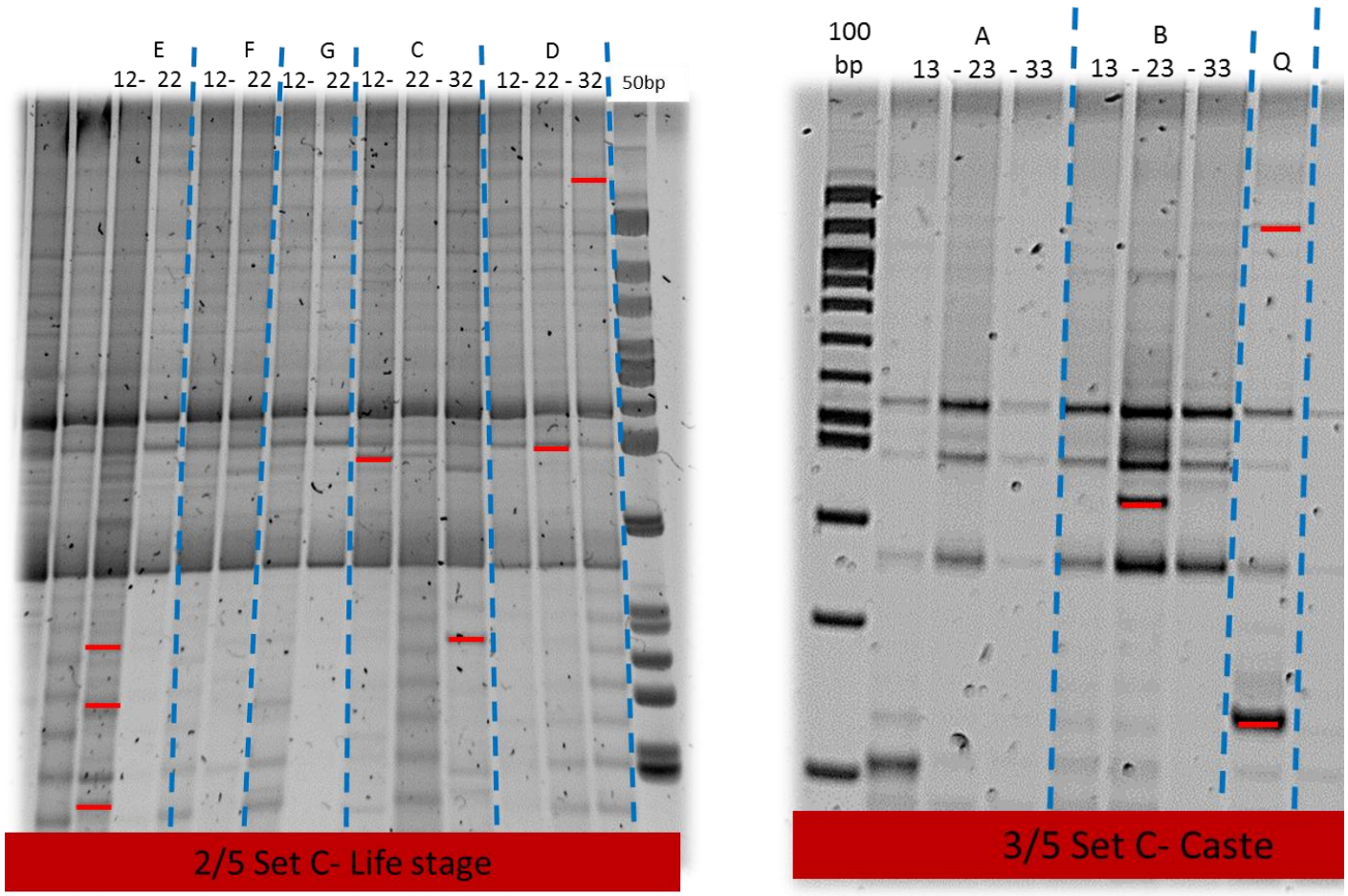


Figure 4.12: 2 AIMS gels for replicate 2 and 3, using primer set C.

C= callow workers, D= pupae, E=L4 larvae, F= L3 larvae and G= L2 larvae

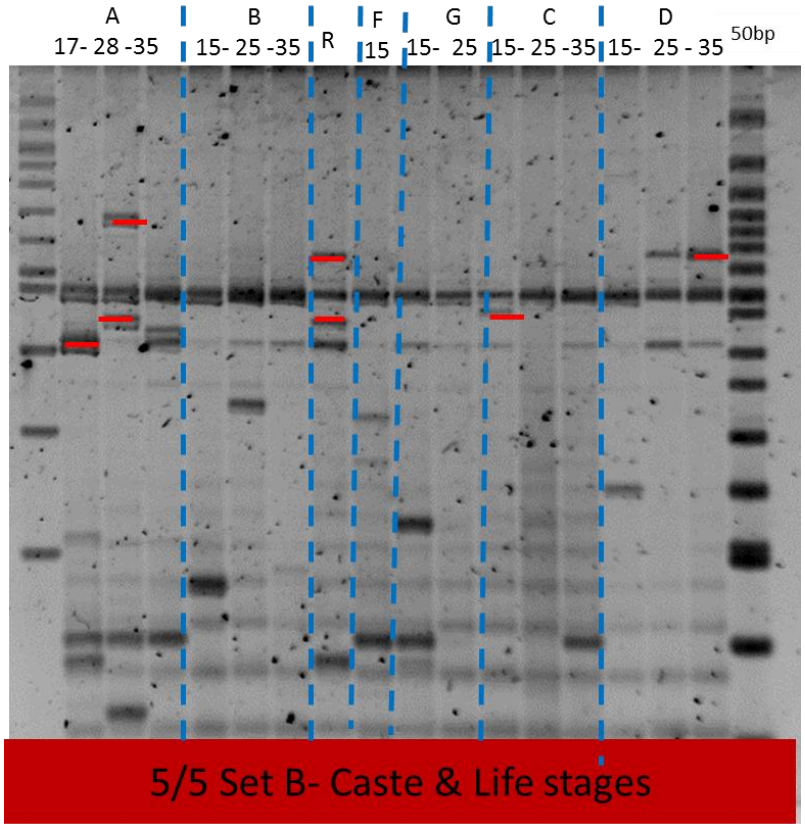


Figure A13: AIMS for replicate 5 with primer set B.

A=10 day old reproducing workers, B= 10 day old non-reproducing workers, r= queen-right reproducing workers and Q= queen

C= callow workers, D= pupae, E=L4 larvae, F= L3 larvae and G= L2 larvae

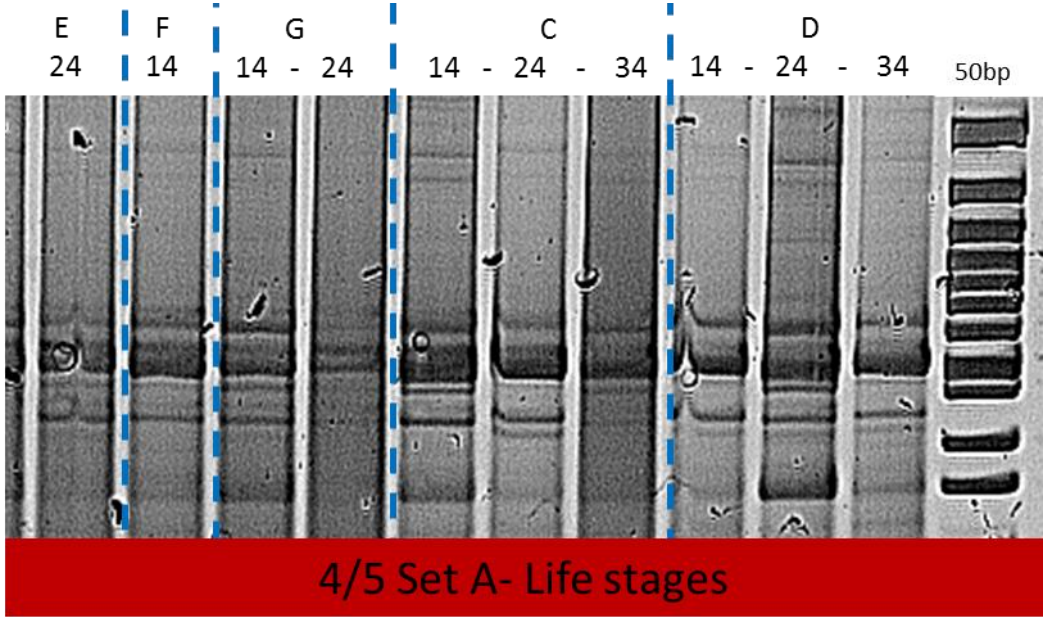


Figure A14: AIMS for replicate 4 with primer set A.

C= callow workers, D= pupae, E=L4 larvae, F= L3 larvae and G= L2 larvae

References

Alaux, C., Hefetz, A. & Jaisson, P. (2006b) Plasticity of worker reproductive strategies in *Bombus terrestris*: lessons from artificial mixed-species colonies. *Animal Behaviour*, **62**, 1417 – 1425.

Alaux, C., Boutot, M., Jaisson, P. & Hefetz, A. (2007) Reproductive plasticity in bumblebee workers (*Bombus terrestris*) – reversion from fertility to sterility under queen influence. *Behavioural Ecology and Sociobiology*, **62**, 213 – 222.

Barchuk, A., Bitondi, M. & Simoes, Z. (2002) Effects of juvenile hormone and ecdysone on the timing of vitellogenin appearance in heamolymph of queen and working workers pupae of *Apis mellifera*. *Journal of Insect Science*, **2**, 1.

Beekman, M. & Oldroyd, B. P. (2008) When Workers Disunite: Intraspecific Parasitism by Eusocial Bees. *Annual Review of Entomology*, **53**, 19 – 37.

Bestor, T.H. (1992) Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. *European Molecular Biology Organization Journal*, **11**, 2611-2617.

Bigot, Y., Jegot, G., Casteret, S., Aupinel, P. & Tasel, J.-N. (2010) DNA modifications & genome rearrangements during the development and sex differentiation of the bumble bee *Bombus terrestris*. *Insect Molecular Biology*. **20**, 165 – 175.

Bird, A. (1995*)* Gene number, noise reduction and biological complexity. *Trends in Genetics*, **11***,* 94–100.

Bird, A. (2002*)* DNA methylation patterns and epigenetic memory. *Genes & Development*, **16***,* 6–21.

Bossdorf, O., Richards, C. & Pigliucci, M. (2008) Epigenetics for ecologists. *Ecology Letters*, **11**, 106-115.

Bourke, A. F. G. & Ratnieks, F. L. W. (2001) Kin-selected conflict in the bumblebee *Bombus terrestris* (Hymenoptera: Apidae). *Proceedings of the Royal Society London* B, **268**, 347 – 355.

Bourke, A. F. G. (1988) Worker reproduction in higher Eusocial hymenoptera. *Quarterly Review of Biology*, **63**, 451 – 461.

Bourke, A. F. G. (1994) Worker matricide in social bees and wasps. *Journal of Theoretical Biology*, **167**, 283 – 292.

Bowler, P. (1983) The Eclipse of Darwinism. Baltimore, Johns Hopkins University Press.

Brown, T. (2001) Southern Blotting and Related DNA Detection Techniques. *Encyclopedia of Life Sciences*, Nature Publishing Group.

Burt, A. & Trivers, R. (2006) Genes in conflict. Harvard University Press, Cambridge, MA.

van Buskirk, J. & Steiner, U. (2009) The fitness costs of developmental canalization and plasticity. *Journal of Evolutionary Biology*, **22**, 852–860.

Callebaut, I., Courvalin J.C. & Mornon J.P. (1999) The BAH (bromo-adjacent homology) domain: a link between DNA methylation, replication and transcriptional regulation. *Federation of European Biochemical Societies Letters*, **446**, 189–93.

Cardoen, D., Wenseleers, T., Ernst, U. R., Danneels, E. L., Laget, D., de Graaf, D. C., Schoofs, L. & Verleyen, P. (2011) Genome-wide analysis of alternative reproductive phenotypes in honeybee workers. *Molecular Ecology*, **20**, 4070 – 4084.

Cassidy S. & Schwartz, S. (1998) Prader-Willi and Angelman Syndromes: Disorders of Genomic Imprinting. *Medicine*, **77**, 140–151.

Chen, T., Tsujimoto, N. & Li, E. (2004) The PWWP domain of Dnmt3a and Dnmt3b Is required for directing DNA methylation to the major satellite repeats at pericentric heterochromatin. *Molecular & Cellular Biology*, **24**, 9048-58

Chittka, A. & Chittka, L. (2010) Epigenetics of Royalty. *PLoS Biology*, **8**.

Ciudad, L., Piulachs, M. D. & Bellés, X. (2006) Systemic RNAi of the cockroach vitellogenin receptor results in a phenotype similar to that of the *Drosophila* yolkless mutant. *FEBS Journal*, **273**, 325 – 335.

Clayton, C. (2007) Methylation & sequence analysis of detoxification gene promoter sequences in *Anopheles gambiae*. MSc thesis, Liverpool School of Tropical Medicine, University of Liverpool.

Cnaani, J., Borst, D. W., Huang, Z. Y., Robinson, G. E. & Hefetz, A. (1997) Caste determination in *Bombus terrestris*: differences in development and rates of JH biosynthesis between queen and worker larvae. *Journal of Insect Physiology*, **43**, 373-381.

Cnaani, J., Schmid-Hempel, R. & Schmidt, J. O. (2002) Colony development, larval development and worker reproduction in *Bombus impatiens* Cresson. *Insectes Sociaux*, **49**, 164 – 170.

Cook, J. M. & Crozier, R. H. (1995) Sex determination and population biology in the Hymenoptera. *Trends in Ecology and Evolution*, **10**, 281 – 286.

Crespi, B. & Yanega, D. (1995) The definition of eusociality. *Behavioural Ecology*, **6**, 109–115.

Crews, D. (2008) Epigenetics and its implications for behavioural neuroendocrinology. *Frontiers in Neuroendocrinology*, **29**, 344−357.

Dong, A., Yoder, J., Zhang, X., Zhou, L., Bestor, T. & Cheng, X. (2001) Structure of human DNMT2, an enigmatic DNA methyltransferase homolog that displays denaturant-resistant binding to DNA. *Nucleic Acids Research.* **29**, 439-448.

Drapeau, M, Albert, S., Kucharski, R., Prusko, C. & Maleszka, R. (2006). Evolution of the Yellow/Major Royal Jelly Protein family and the emergence of social behaviour in honey bees. *Genome Research*, **16**, 1385-1394.

Duchateau, M. & Velthuis, H. (1988) Development and reproductive strategies in *Bombus terrestris* colonies. *Behaviour*, **107**, 186-20.

Duchateau, M. & Velthuis, H. H. W. (1989) Ovarian development and egg laying in workers of *Bombus terrestris*. *Entomologia Experimentalis et Applicata*, **51**, 199 – 213.

Elango, N., Hunt, B. G., Goodisman, M. A. D. & Soojin, V. Y. (2009) DNA methylation is widespread and associated with differential gene expression in castes of honeybee, *Apis mellifera. Proceedings of the National Academy of Sciences of the United States*, **106**, 11206 – 11211.

Ernande, B. & Dieckmann, U. (2004) The evolution of phenotypic plasticity in spatially structured environments: implications of intraspecific competition, plasticity costs and environmental characteristics. *Journal of Evolutionary Biology*, **17**, 613–628.

Espinosa-Soto, C., Martin, O., & Wagner, A. (2011) Phenotypic plasticity can facilitate adaptive evolution in gene regulatory circuits. *BMC Evolutionary Biology*, **11**, (1), 5.

Field, L. M., Blackman, R. L., Tyler-Smith, C. & Devonshire, A.L. (1999*)* Relationship between amount of esterase and gene copy number in insecticide-resistant *Myzus persicae* (Sulzer). *Biochemical Journal*, **339**, 737–742.

Field, L. M. (2000*)* Methylation and expression of amplified esterase genes in the aphid *Myzus persicae* (Sulzer). *Biochemical Journal,* **349**, 863–868.

Field, L. M., Lyko, F., Mandrioli, M. & Prantera, G. (2004) DNA methylation in insects. *Insect Molecular Biology,* **13**, 109-115.

Finnegan, E.J., Peacock, W.J. & Dennis, E.S. (1996*)* Reduced DNA methylation in *Arabidopsis thaliana* results in abnormal plant development. *Proceedings of the National Academy of Sciences of the United States,* **93**, 8449–8454.

Finnegan, E. (2002) Epialleles—a source of random variation in times of stress. *Current Opinion in Plant Biology*, **5**, 101−106.

Fisher, O., R. Siman-Tov, A. & Ankri, S. (2004) Characterization of cytosine methylated regions and 5-cytosine DNA methyltransferase (Ehmeth) in the protozoan parasite *Entamoeba histolytica. Nucleic Acids Research*, **32**, 287–297.

Foret, S., Kucharshi,R., Pittelkow, Y., Lockette, G. & Maleszeka, R. (2009) Epigenetic regulation of the bee trascriptome: unravelling the nature of methylated genes. *BMC Genomics*, **10**, (472).

Fraga, M.F., Ballestar, E., Paz, M.F., Ropero, S., Setien, F., Ballestart, M.L., Heine-Suner, D., Cigudosa, J., Urioste, M., Benitez, J., Biox-Chornet, M., Sanchez-Aguilera, A., Ling, C., Carlsson, E., Poulsen, P., Vaag,A., Stephen, Z., Spector, T., Wu, Y., Plass, C. & Esteller, M. (2005) Epigenetic differences arise during the lifetime of

monozygotic twins. *Proceedings of the National Academy of Sciences of the United States,* **102**, 10604–10609.

Frigola, J., Ribas, M., Risques, R. & Peinado, M. (2002) Methylome profiling of cancer cells by amplification of inter-methylated sites (AIMS). *Nucleic Acids Research*, **30**.

Fry, J. (2003) Multilocus models of sympatric speciation: Bush versus Rice versus Felsenstein. *Evolution*, **57**, 1735–1746.

Gabor, D. & Maleszeka, R. (2011) Epigenomic communication systems in humans and honey bees: From molecules to behaviour. *Special Issue: Behavioral Epigenetics*, **59**, 399–406.

Gasser, R., Hu, M., Chilton, Campbell, B., Jex, A., Otranto, D., Cafarchia, C., Beveridge, I. & Zhu, X. (2006), Single-strand confirmation polymorphism (SSCP) for the analysis of genetic variation. *Nature Protocols*, **1**, 3121 – 3128.

Gaudet, F., Hodgson, J., Eden, A., Jackson-Grusby, L., Dausman, J., Gray, J.W., Leonhardt, H. & Jaenisch, R. (2003) Induction of tumours in mice by genomic hypomethylation. *Science,* **300**, 489–492.

Goll, M., & Bestor, T. (2005). Eukaryotic cytosine methyltransferases. Annual Review of Biochemistry, **74**, 481-514.

Goulson, D. (2003) Bumblebees – Their Behaviour and Ecology. Oxford University Press, London, U.K.

Goulson, D., Stout, J. C. & Kells, A.R. (2002a) Do alien bumblebees compete with native flower-visiting insects in Tasmania*? Journal of Insect Conservation,* **6**, 179-189.

Gourbiere, S. (2004) How do natural and sexual selection contribute to sympatric speciation? *Journal of Evolutionary Biology*, **17**, 1297–1309.

Graff, J., Jemielity, S., Parker, J. D., Parker, K. M. & Keller, L. (2007) Differential gene expression between adult queens and workers in the ant *Lasius niger*. *Molecular Ecology*, **16**, 675–683.

Grether, G. (2005) Environmental change, phenotypic plasticity, and genetic compensation. *American Naturalist*. **166**, 115–123.

Guidugli, K. R., Piulachs, M., Bellés, X. & Lourenço, A. P. (2005) Vitellogenin expression in queen ovaries and in larvae of both sexes of *Apis mellifera*. *Archives of Insect Biochemistry and Physiology*, **59**, 211–218.

Guidugli-Lazzarini, K.R., Nascimento' A., Tanaka, E.D., Piulachs, M. D., Hartfelder, K., Bitondi, , G.M. & Simões, Z.L.P. (2008) Expression analysis of putative vitellogenin and lipophorin receptors in honey bee (*Apis mellifera* L.) queens and workers. *Journal of Insect Physiology*, **54**, 1138 – 1147.

Guzman-Novoa, E., Hunt, G. J., Page Jr. R. E., Uribe-Rubio, J. L., Prieto-Merlos, D. & Becerra-Guzman, F. (2005) Paternal effects on the defensive mechanism of honeybees. *Journal of Heredity*, **96**, 376 – 380.

Haig, D. & Westoby, M. (1989) Parent-specific gene expression and the triploid endosperm. *American Naturalist*, **134**, 147-155.

Haig, D. (1997) Parental antagonism, relatedness asymmetries, and genomic imprinting. *Proceedings of the Royal Society of London B*, **264**, 1657-1662.

Haig, D. (2000) The kinship theory of genomic imprinting. *Annual Review of Ecology & Systematics*. **31**, 9-32.

Haig, D. (2000b) Genomic imprinting, sex-biased dispersal, and social behaviour. *Annals of the New York Academy of Science*, **907**, 149 – 163.

Haig, D. (2004) Genomic imprinting and kinship: how good is the evidence? *Annual Review of Genetics*, **38**, 553-585.

Haig, D. (2011) Lamarck ascending! A review of transformations of Lamarckism: From Subtle Fluids to Molecular Biology, edited by Gissis, S. & Jablonka, E. (2011) *Philosophy and Theory in Biology*. MIT Press.

Haig, D. (2011). Genomic imprinting and the evolutionary psychology of human kinship. *Proceedings of the National Academy of Sciences*, **108** (Supplement 2), 10878-10885..

Hamilton, W. D. (1964) The genetics evolution of social behaviour (I & II). *Journal of Theoretical Biology*, **7**, 1 – 52.

Harris, R.A., Wang, T., Coarfa, C., Nagarajan, R.P., Hong, C., Downey, S.L., Johnson, B.E., Fouse, S.D., Delaney, A., Zhao, Y., Olshen, A., Ballinger, T., Zhou, X., Forsberg, K., Gu, J., Echipare, L., O'Green, H., Lister, R., Pelizzola, M., Xi, Y., Epstein, C., Bernstein, B., Hawkins, R., Ren, B., Chung, W., Gu, H., Bock, C., Gnirkle, A., Zhang, M., Haussler, D., Ecker, J., Li, W., Farnham, P., Waterland, R., Meissner, A., Marra, M., Hirst, M., Milsosavljevic, A. & Costello, J. (2010). Comparison of sequencing-

based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. *Nature Biotechnology*. **28**, 1097–1105.

Hendrich, B. & Tweedie, S. (2003*)* The methyl-CpG binding domain and the evolving role of DNA methylation in animals. *Trends in Genetics*, **19**, 269–277.

Herrmann, F., Westphal, C., Moritz, R. & Steffan-Dewenter, I. (2007). Genetic diversity and mass resources promote colony size and forager densities of a social bee (*Bombus pascuorum*) in agricultural landscapes. *Molecular Ecology*, **16,** (6), 1167-1178.

Hick, C.A., Field, L.M. & Devonshire, A.L. *(*1996*)* Changes in the methylation of amplified esterase DNA during loss and reselection of insecticide resistance in peach-potato aphids, *Myzus persicae. Insect Biochemistry and Molecular Biology*, **26**, 41–47.

Hoffman, E. A. & Goodisman, M. A. D. (2007) Gene expression and the evolution of phenotypic diversity in social wasps. *Bio Medical Central Biology*, **5**, 23.

Holsboer, F., Wotjak, C., Almeida, O. & Spengler, D. (2010) Dynamic DNA methylation programs persistent adverse effects of early-life stress. *Nature Neuroscience,* **5**.

Hongyo, T., Buzzard, G., S., Calvert, R. J. & Weghorst, C. M. (1993) Cold SSCP- a simple, rapid and non-radioactive method for optimised single strand confirmation polymorphism analysis. *Nucleic Acids Research*, **21**, 3637-3642.

Hughes, K., Burleson, M. & Rodd, F. (2002) Is phenotypic plasticity adaptive? In: Biodemography of Fertility. Edited by Rodgers, J. & Kohler, H. Kluwer Academic Press, Dordrecht, the Netherlands, 23-42.

Hurst, L. D., 2009. Evolutionary genomics and the reach of selection. *Journal of Biology*, 8, 12.

Jackson-Grusby, L., Beard, C., Possemato, R., Tudor, M., Fambrough, D., Csankovszki, G., Dausman, J., Lee, P., Wilson, C., Lander, E. & Jaenisch, R. (2001*)* Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. *Nature Genetics*, **27***,* 31–39.

Jaenisch, R. (1997) DNA methylation and imprinting: why bother? *Trends in Genetics*, **13**, 323-329.

Jeltsch,A. (2002) Beyond Watson & Crick: DNA methylation & molecular enzymology of DNA methyltransferases. *ChemBioChem.* **3**, 274-293.

Jorda, M., Rodriguez, J., Frigola, J. & Peinado, M. (2009) Analysis of DNA methylation by amplification of intermethylated sites (AIMS). *Methods in Molecular Biology*, **507**, 107-116.

Kass, S. U., Pruss, D. & Wolffe, A. P. (1997) How does DNA methylation repress transcription? *Trends in Genetics*, **13**, 444-449.

Kelly, M., Sanford, E., & Grosberg, K. (2012). Limited potential for adaptation to climate change in a broadly distributed marine crustacean. *Proceedings of the Royal Society B: Biological Sciences*, **279**, (1727), 349-356.

Kelly,S., Panhuis, T. & Stoehr, A. (2012) Phenotypic Plasticity: Molecular Mechanisms and Adaptive Significance. *Comprehensive Physiology*, 1417–1439.

Keyte, A., Percifield, R., Liu, B. & Wendel, J. (2006) Infraspecific DNA methylation polymorphism in cotton (*Gossypium hirsutum* L.). *Journal of Heredity*, **97**, 444−450.

Komori, H., LaMere, S., Torkamani, A., Hart, G., Kotsopoulos, S., Warner, J., Samuels, M., Olson, J., Head, S., Ordoukhanian, P., Lee, P., Link, D. & Salomon, D. R. (2011) Application of microdroplet PCR for large-scale targeted bisulfite sequencing. *Genome Research*, **21** (10), 1738-1745.

Khosla, S., Mendiratta, G. & Brahmachari, V. (2006) Genomic imprinting in the mealybugs. *Cytogenetic and Genome Research*, **113**, 41-52.

Koywiwattrakul, P., Thompson, G.J., Sitthipraneed, S., Oldroyd, B.P. & Maleszka, R. (2005) Effects of carbon dioxide narcosis on ovary activation and gene expression in worker honeybees, *Apis mellifera. Journal of Insect Science*, **5**.

Kroll, J. (2007) Molecular Chaperones and the Epigenetics of Longevity and Cancer Resistance. *Annals of the New York Academy of Sciences*, **1100**, 75-83.

Kronauer, D. J. C. (2008) Genomic imprinting & kinship in the social Hymenoptera: What are the predictions? *Journal of Theoretical Biology*, **254**, 737– 740.

Kronforst, M. R., Gilley, D.C., Strassmann, J. E., Queller, D. C. (2008) DNA methylation is widespread across social Hymenoptera. *Current Biology*, **18**, R287–R288.

Kucharski, R., Maleszeka, J., Foret, S. & Malkeszka, R. (2008) Nutritional control of reproductive status in honeybees via DNA methylation. *Science*, **319**, 1827 – 1830.

Kunert, N., Marhold, J., Stanke, J., Stach, D. & Lyko, F. *(*2003*)* A Dnmt2-like protein mediates DNA methylation in *Drosophila. Development*, **130**, 5083–5090.

Legeali, F., Rizk, G., Walsh, T., Edwards, O., Gordon, C., Lavenier, D., Leterme,N., Mereau, A., Nicolas, J., Tagu, D. & Jaubert-Possamai, S. (2010) Bioinformatic prediction, deep sequencing of microRNAs and expression analysis during phenotypic plasticity in the pea aphid, *Acyrthosiphon pisum*. *BMC Genomics*, **11**, 281-280.

Lei, H., Oh, S., Okano, M., Juttermann, R., Goss, K., Jaenisch K. & Li, E. (1996) *De novo* DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development*, **10**, 3195-3205.

Levy, S. & Siegal, M. (2008) Network hubs buffer environmental variation in *Saccharomyces cerevisiae*. *PloS Biology*, **6**, e264.

Lister, R. & Ecker, J.R. (2009) Finding the fifth base: Genome-wide sequencing of cytosine methylation. *Genome Research*, **19**, 959-66.

Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J., Lee, L., Ye, Z., Ngo, Q., Edsall, L., Antosiewicz-Bourget, J., Stewart, R., Ruotti, V., Millar, A., Thomson, J., Ren, B. & Ecker, J. (2009) Human DNA

methylomes at base resolution show widespread epigenomic differences. *Nature*, **462**, 315–322.

Liu, K., Wang, Y. F., Cantemir, C. & Muller, M. T. (2003). Endogenous assays of DNA methyltransferases: Evidence for differential activities of *DNMT1*, *DNMT2*, and *DNMT3* in mammalian cells *in vivo*. *Molecular and Cellular. Biology*, **23**, 2709-2719.

Lizé, A., Cortecero, A.M. , Atlan, A. & Poinsot, D . (2007) Kin recognition in *Aleochara bilineata* could support the kinship theory of genomic imprinting. *Genetics*, **175**, 1735 – 1740.

Lockett, G. A., Helliwell, P. & Maleszeka, R. (2010) Involvement of DNA methylation in memory processing in the honey bee. *Neuroreport*, **21**, 812 – 816.

Luedi, P. Dietrich, F., Weidman, J., Bosko, J., Jirtle, R. & Hartemink, A. (2007) Computational and experimental identification of novel human imprinted genes. *Genome Research*, **17**, 1723-1730.

Lyko, F. (2001) DNA methylation learns to fly. *Trends in Genetics*, **17**, 169-172.

Lyko, F., Ramsahoye, B.H. & Jaenisch, R. (2000*)* DNA methylation in *Drosophila melanogaster*. *Nature*, 408*,* 538–540.

Lyko, F., Foret, S., Kucharski, R., Wolf, S., Falckenhayn, C. & Maleszka, R. (2010) The honey bee epigenomes: differential methylation of brain DNA in queens and workers. *PLoS Biol* 8: e1000506.

Lyko, F. & Maleszeka, R. (2011) Insects as innovative models for functional studies of DNA methylation. *Trends in Genetics*, **27**, 127 – 131.

Ma, P., Lin, S., Bartolomei, M. & Schultz, R. (2010) Metastasis tumor antigen 2 (MTA2) is involved in proper imprinted expression of *H19* and *Peg3* during mouse pre-implantation development. *Biology of Reproduction*, **83**, 1027-1035.

Mandrioli, M. & Volpi, N. *(*2003*)* The genome of the lepidopteran *Mamestra brassicae* has a vertebrate-like content of methyl-cytosine. *Genetica*, **119**, 187–191.

Mares, S., Ash, L. & Gronenberg, W. (2005) Brain allometry in bumblebee and honey bee workers. *Brain, Behaviour and Evolution*, **66**, 50 – 61.

Marhold, J., Rothe, N., Pauli, A., Mund, C., Kuehle, K., Brueckner, B. & Lyko, F. (2004) Conservation of DNA methylation in dipteran insects. *Insect Molecular Biology,* **13**, 117–124.

Martin, S .J., Beekman, M., Wossler, T. C., Ratnieks, F. L. W. (2002b) Parasitic Cape honeybee workers, *Apis mellifera capensis*, evade policing. *Nature*, **415**, 163 – 165.

Maynard-Smith, J. & Szathmáry, E. (1995) The major transitions in evolution. Oxford: W. H. Freeman.

Moran, N. (1992) The evolutionary maintenance of alternative phenotypes. *American Naturalist*, **139**, 971–989.

Morozova, O., Hirst, M. & Marra, A. (2009) Applications of new sequencing technologies for transcriptome analysis. *Annual Review* of *Genomics* and *Human Genetics*, **10**, 135–151.

Murgatroyd, C., Patchev, A., Wu, Y., Micale, V., Bockmühl, Y., Fischer, D., Murphy, S. & Jirtle, R. (2003) Imprinting evolution and the price of silence. *BioEssays* **25**, 577–588.

Nan, X., Meehan, R. R., & Bird, A. (1993) Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. *Nucleic Acids Research*, **21**, 4886-4892.

Nanork, P., Chapman, N., Wongsiri, S., Lim, J., Gloag, R. & Oldroyd, B. (2006) Social parasitism by workers in queen-less and queen-right *Apis cerana* colonies. *Molecular Ecology*, **16**, 1107 – 1114.

Nelson, C. M., Ihle, K. E., Fondrk, M. K., Page, R. E. & Amdam, G. (2007) The gene vitellogenin has multiple coordinating effects on social organization. *PLoS Biology*, **5**, 673 – 677.

Newman, R. (1992) Adaptive plasticity in amphibian metamorphosis. What type of phenotypic variation is adaptive and what are the costs of such plasticity? *Bioscience*, **42**, (9), 671-678.

Nielsen, D., Novoradovsky, A. & Goldman, D. (1995) SSCP primer design based on single-strand DNA structure predicted by a DNA folding program. *Nucleic Acids Research*, **23**, 2287–2291.

Olsson, M. & Shine, R. (1999). Plasticity in frequency of reproduction in an alpine lizard, *Niveoscincus microlepidotus*. *Copeia*, 794-796.

Ono, M., Swanson, J.J., Field, L.M., Devonshire, A.L. & Siegfried, B.D. (1999) Amplification and methylation of an esterase gene associated with insecticide-resistance in greenbugs, *Schizaphis graminum* (Rondani) (Homoptera: Aphididae). *Insect Biochemistry and Molecular Biology*, **29**, 1065–1073.

Panning, B. & Jaenisch, R. (1996) DNA hypomethylation can activate *Xist* expression and silence X-linked genes. *Genes & Development*, **10**, 1991–2002.

Pennings, S., Allan, J. & Davey, C. S. (2005) DNA methylation, nucleosome formation and positioning. *Briefings in Functional Genomics & Proteomics*, **3**, 351–361.

Pereboom, J. J. M., Jordan, W. C., Sumner, S., Hammond, R. L. & Bourke, A. F. G. (2005) Differential gene expression in queen-worker caste determination in bumble-bees. *Proceedings of the Royal Society B*, **272**, 1145 – 1152.

Pigllucci, M. (1996) How organisms respond to environmental changes: from phenotypes to molecules (and vice versa). *Trends in Ecology & Evolution*, **11**, 168-173.

Pigliucci, M. (2001) Phenotypic Plasticity: Beyond Nature and Nurture. Johns Hopkins University Press, Baltimore.

Qiu, C., Sawada, K., Zhang, X. & Cheng, X. (2002) The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds. *Nature Structural Biology*, **9**, 217-24.

Queller, D. C. & Strassmann, J. E. (2002) The many selves of social insects. *Science*, **296**, 311 – 313.

Queller, D.C. (2003) Theory of genomic imprinting conflict in social insects. *Bio Medical Central: Evolutionary Biology*, **3**, 15.

Ramsahoye, B.H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A.P. & Jaenisch, R. *(*2000*)* Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proceedings of the National Academy of Sciences of the United States*, **97**, 5237–5242.

Ratnieks, F. L. W. & Wenseleers, T. (2007) Altruism in insect societies and beyond: voluntary or enforced? *TRENDS in Ecology and Evolution*, **23**, 45 – 52.

Ratnieks, F. L. W. (1988) Reproductive harmony via mutual policing by workers in eusocial Hymenoptera. *American Naturalist*, **132**, 217 – 236.

Ratnieks, F. L. W., Foster, K. R., & Wenseleers, T. (2006) Conflict resolution in insect societies. *Annual Review of Entomology*. **51**, 581 – 608.

Razin A (1998) CpG methylation, chromatin structure and gene silencing—a three-way connection. *EMBO J*, **17**, 4905–4908.

Razin, A. & Cedar, H. *(*1994*)* DNA methylation and genomic imprinting. *Cell*, **77**, 473–476.

Reik, W., Dean, W. & Walter, J. (2001) Epigenetic reprogramming in mammalian development. *Science*, **293,** (5532), 1089–93.

Reinders, J., Wulff, B. B., Mirouze, M., Marí-Ordóñez, A., Dapp, M., Rozhon, W., Bucher, E., Theiler, G. & Paszkowski, J. (2009). Compromised stability of DNA methylation and transposon immobilization in mosaic Arabidopsis epigenomes. *Genes & development*, **23**, (8), 939-950.

Richards, C., Bossdorf, O. & Pigliucci, M. (2010) What Role Does Heritable Epigenetic Variation Play in Phenotypic Evolution? *BioScience*, **60**, 232–237.

Riddell, C., Adams, S., Schmid-Hempel, P. & Mallon, E. (2009) Differential expression of Immune Defences is associated with specific host-parasite interactions in insects. *Plos One*, **10,** 135-138 e7621.

Riddel C. E., Sumner, S., Adams, S., & Mallon, E. B. (2011) Pathways to immunity: temporal dynamics of the bumblebee (*Bombus terrestris*) immune response against a trypanosome gut parasite. *Insect Molecular Biology*, **20**, 529-540.

Riek, W. & Walter, J. (2001) Genomic imprinting: parental influence on the genome. *Nature Reviews Genetics*, **2**, 21-32.

Robinson, K.L., Tohidi-Esfahani, D., Lo, N., Simpson, S. & Sword, G. (2011) Evidence for widespread genomic methylation in the migratory locust, *Locusta migratoria* (Orthoptera: Acrididae). *PLOS One*, **6**, 12.

Rountree, M. R., Bachman, K. E. & Baylin, S. B. (2000) *Dnmt1* binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nature Genetics*, **25**, 269-277.

Ruden, D., Garfinkel, M., Xiao L., Lu, X. (2005) Epigenetic regulation of trinucleotide repeat expansions and contractions and the "biased embryos" hypothesis for rapid morphological evolution. *Current Genomics*, **6**, 145−155.

Saitou, N., & Nei, M. (1987) The neighbour-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, **4**, (4), 406-425.

Salmon, A., Clotault, J., Jenczewski, E., Chable, V. & Manzanares-Dauleux, M. (2008) *Brassica oleracea* displays a high level of DNA methylation polymorphism. *Plant Science*, **174**, 61−70.

Sambrook, J., MacCallum, P. & Russell, D. (2000) Molecular Cloning: A Laboratory Manual (Third Edition).

Santoni, V., Bellini, C., & Caboche, M. (1994) Use of two-dimensional protein-pattern analysis for the characterization of *Arabidopsis thaliana* mutants. *Planta*, **192**, (4), 557-566.

Schaefer, M. & Lyko, F. (2007) DNA methylation with a sting: an active DNA methylations system with the honeybee. *BioEssays*, **29**, 208 – 211.

Schaefer, M., & Lyko, F. (2010). Solving the *Dnmt2* enigma. *Chromosoma*, **119**, (1), 35-40.

Schmid-Hempel, R. & Schmid-Hempel, P. (2000) Female mating frequencies in *Bombus* spp. from central Europe. *Insectes Sociaux*, **41**, 36 – 41.

Shimizu, T.S., Takahashi, K. & Tomita, M. *(*1997*)* CpG distribution patterns in methylated and non-methylated species. *Gene,* **205**, 103–107.

Sivashankari, S. & Shanmughavel, P. (2006) Functional annotation of hypothetical proteins – A review. *Bioinformation*, **1** (8), 335–338.

Song, L., James, S., Kazim, L. & Karpf, A. (2005) Specific method for the determination of genomic DNA methylation by liquid chromatography-electrospray ionization tandem mass spectrometry. *Analytical Chemistry*, 504-510.

Soppe, W.J., Jasencakova, Z., Houben, A., Kakutani, T., Meister, A., Huang, M.S., Jacobsen, S.E., Schubert, I. & Fransz, P.F. *(*2002*)* DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in *Arabidopsis. European Molecular Biology Organization Journal*, **21**, *6549–6559.*

Spielman, M., Vinkenoog, R., Dickinson, H. G. & Scott, R. J. (2001) The epigenetic basis of gender in flowering plants and mammals. *Trends in Genetics,* **17**, 705–711.

Spitze, K. (1992). Predator-mediated plasticity of prey life history and morphology: *Chaoborus americanus* predation on *Daphnia pulex*. *American Naturalist*, 229-247.

Stancheva, I. & Meehan, R.R. *(*2000*)* Transient depletion of *Dnmt1* leads to premature gene activation in *Xenopus* embryos. *Genes & Development*, **14**, 313–327.

Stancheva, I., Hensey, C. & Meehan, R.R. *(*2001*)* Loss of the maintenance methyltransferase, *Dnmt1*, induces apoptosis in *Xenopus* embryos. *European Molecular Biology Organization Journal*, **20**, 1963–1973.

Stolle, E., Wilfert, L., Schmid-Hempel, R., Schmid-Hempel, P., Kube, M., Reinhardt, R. & Moritz, R. F. A. (2011) A second generation genetic map of the bumblebee *Bombus terrestris* (Linnaeus, 1758) reveals slow genome and chromosome evolution in the Apidae. *BMC Genomics*, **12.**

Storz, J., Scott, G., & Cheviron, Z. A. (2010) Phenotypic plasticity and genetic adaptation to high-altitude hypoxia in vertebrates. *The Journal of Experimental Biology*, **213**, (24), 4125-4136.

Sultan, S. & Spencer, H. (2002) Metapopulation structure favours plasticity over local adaptation. *American Naturalist*, **160**, 271–283.

Sumner. S., Pereboom, J.J.M. & Jordan W.C. (2006). Differential gene expression and phenotypic plasticity in behavioural castes of the primitively eusocial wasp, *Polistes canadensis*. *Proceedings of the Royal Society of London, Series B*, **273**, 19-26.

Suzuki, M. & Bird, A. (2008) DNA methylation landscapes: provocative insights from epigenomics. *Nature Reviews Genetics*, **9**, 465 – 476.

Thibert-Plante, X. & Hendry, A. P. (2011) The consequences of phenotypic plasticity for ecological speciation. *Journal of Evolutionary Biology*, **24**, (2), 326-342.

Toth, A. L., Varala, K., Newman, T., Miguez, F., Hutchison, S., Willoughby, D., Sion, J., Egholm, M., Hunt, J., Hudson, M. & Robinson, G. (2007) Wasp gene expression supports an evolutionary link between maternal behaviour and eusociality. *Science*, **318**, 440 – 444.

Tweedie, S., Charlton, J., Clark, V. & Bird, A. *(*1997*)* Methylation of genomes and genes at the invertebrate–vertebrate boundary. *Molecular & Cell Biology*, **17**, 1469–1475.

Ullmann, A., Jacob, F. & Monod, J. (1967) Characterization by in vitro complementation of a peptide corresponding to an operator-proximal segment of the beta-galactosidase structural gene of *Escherichia coli*. *Journal of Molecular Biology*, **24** (2), 339–343.

Varmuza, S. & Mann, M. (1994) Genomic imprinting-defusing the time bomb theory. *Trends in Genetics*, **10**, 118-123.

Walsh, C.P. & Bestor, T.H. *(*1999*) Cytosine* methylation and mammalian development. *Genes & Development*, **13**, 26–34.

Walsh, T., Brisson, J. A., Robertson, H. M., Gordon, K., Jaubert-Possamai, S., Tagu, D. & Edwards, O.R. (2010) A functional DNA methylation system in the pea aphid *Acyrthosiphon* pisum. Insect *Molecular Biology*, **19**, 215 – 228.

Wang, Y., Jorda, M., Jones, P., Maleszeka, R., Ling, X., Robertson, H., Mizzen, C., Peinado, M. & Robinson, G. (2006) Functional CpG methylation system in a social insect. *Science*, **314**, 645 – 647.

Wang, Y., Kocher, S., Linksvayer, T.A., Grozinger, C. M., Page, R. & Amdam, G. (2012) Regulation of behaviourally associated gene networks in worker honey bee ovaries. *Journal of Experimental Biology*, **215**, 124-134.

Wang,Y. & Leung, F. C. C. (2009) *In silico* prediction of two classes of honeybee genes with CpG deficiency or CpG enrichment and sorting according to gene ontology classes. *Journal of Molecular Evolution*. **68**, 700 – 705.

Warren, J. H., Richards, S., Desjardins, C. A., Niehuis, O., Gadau, J. & The Nasonia Genome Working Group (2010) Functional and evolutionary insights from the Genomes of Three Parasitoid *Nasonia* Species. *Science,* **327**, 343-348.

Weismann, A. (1891) The significance of sexual reproduction in the theory of natural selection. In: Poulton, E., Schönland, S. & Shipley, A. (1987) Essays upon Heredity and Kindred Biological Problems. Oxford: Clarendon Press.

Wenseleers, T. & Ratnieks, F. L. W. (2006) Comparative Analysis of Worker Reproduction and Policing in Eusocial Hymenoptera Supports Relatedness Theory. *The American Naturalist*, **168**, E163 – E179.

Wenseleers, T., Helantera, H., Hart, A. & Ratnieks, F. L. W. (2004) Worker reproduction and policing in insect societies: an ESS analysis. *Journal of Evolutionary Biology*, **17**, 1035 – 1047.

Wenseleers, T., Ratnieks, F. L. W., Ribeiro, M., Alves, D. & Imperatriz-Fonseca (2004b) Working class royalty: bees beat the caste system. *Biology Letters*, **1**, 125-128.

West-Eberhard, M. (2005) Developmental plasticity and the origin of species differences. *Proceedings of the National Academy of Science*, **102**, 6543-6549.

Wilfert, L., Gadau, J. & Schmid-Hempel, P. (2006) A core linkage map of the bumblebee *Bombus terrestris*. *Genome*, **49**, 1215 – 1226.

Wilfert, L., Torres, M. M., Reber-Funk, C., Schmid-Hempel, R., Tomkins, J., Gadau, J. & Schmid-Hempel, P. (2009) Construction and characterisation of a BAC-library for a key pollinator, the bumblebee *Bombus terrestris* L. *Insectes Sociaux*, **56**, 44 – 48.

Willmer, P., Bataw, A., & Hughes, J. (2008). The superiority of bumblebees to honeybees as pollinators: insect visits to raspberry flowers. *Ecological Entomology*, **19,** (3), 271-284.

Wilson, E. (1971) The insect societies. Belknap Press of Harvard University Press. Cambridge. Massachusetts.

Wilson, E. O. (2008) One Giant Leap: How Insects Achieved Altruism and Colonial Life. *BioScience*, **58**, 17 – 25.

Wolffe, A.P., Jones, P.L. & Wade, P.A. *(*1999*)* DNA methylation. *Proceedings of the National Academy of Sciences of the United States*, **96**, 5894–5896.

Yoder, J.A., Walsh, C.P. & Bestor, T.H. *(*1997*)* Cytosine methylation and the ecology of intragenomic parasites. *Trends in Genetics*, **13**, 335–340.

Zhao , y., Sun, H., Lu, J., Xiaoxue, J., Chen, X., Tao, D., Huang, W. & Huang, B. (2005) Lifespan extension and elevated *hsp* gene expression in *Drosophila* caused by histone deacetylase inhibitors. *Journal of Experimental Biology*, **208**, 697-705.